

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 11:32:54 ; Search time 29.74 Seconds
(without alignments)
1000.934 Million cell updates/sec

Title: US-09-674-235-1

Sequence: 1406
1 MAFERLSQVGLGRFQMLH.....RKISRKRHKNDYKVKTKF 268

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
21: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*
22: /SIDSL1/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1406	100.0	268	21	AAV52386	Human transmembran
2	1238	88.1	541	22	AAAB47276	HOAT5. Homo sapie
3	1237	88.0	391	22	ABBI1732	Human secreted pro
4	993	70.6	213	22	AAAM24188	Human EST encoded
5	799.5	56.9	553	22	AAE04897	Human transporter
6	768	54.6	554	22	AAAB47275	HOAT4. Homo sapie
7	765	54.4	341	22	AAAM40059	Human polypeptide
8	759	54.0	339	21	AAAB19406	Amino acid sequenc
9	601	42.7	553	22	AAE10332	Human transporter
10	594.5	42.3	369	22	AAAM41845	Human polypeptide
11	592	42.1	550	22	AAE06612	Human protein havi

12	592	42.1	550	22	AAAB69091	Human organic anio
13	590.5	42.0	566	22	AAE10336	Human transporter
14	588.5	41.9	578	22	AAE06571	Human protein havi
15	472	33.6	645	22	AAU32379	Novel human secret
16	458.5	32.6	550	21	AAI44278	Human organic anio
17	458.5	32.6	550	22	AAAB47271	HOAT1. Homo sapie
18	458.5	32.6	553	20	AAAM88489	Human organic anio
19	451.5	32.1	551	20	AAAM88488	Rat organic anion
20	436.5	31.0	545	22	AAAB36553	Mouse organic anio
21	428	30.4	815	22	ABG26899	Novel human diagno
22	390.5	27.8	542	22	AAAB7274	HOAT3. Homo sapie
23	386.5	27.5	542	22	AAAB7274	Novel human secret
24	373	26.5	128	22	AAU3066	Mouse osteoclast t
25	367.5	26.1	537	18	AAAM4195	Rat cerebral organ
26	362.5	25.8	536	21	AAAY92903	Peptide #11836 enco
27	357.5	25.4	561	18	AAAM44330	Protein #9189 enco
28	344	24.5	113	22	ABAB7190	Human brain expres
29	344	24.5	113	22	ABAB7190	Human bone marrow
30	344	24.5	113	22	AAAM5378	Peptide #8388 enco
31	344	24.5	113	22	AAAM8066	Peptide #12329 enc
32	344	24.5	113	22	AAAM21954	Rat liver anion tr
33	344	24.5	113	22	AAAM38292	Murine organic an
34	324.5	23.1	535	21	AAAY51249	HOAT2B. Homo sapi
35	308	21.9	540	22	AAAB9401	HOAT2A. Homo sapi
36	307.5	21.9	538	22	AAAB47272	A human organic an
37	307.5	21.9	546	22	AAAB08823	Human liver specif
38	296.5	21.1	548	21	AAAB12030	A human organic an
39	296.5	21.1	552	22	AAAB12030	Novel human diagno
40	295	21.0	551	21	AAAB08824	Human reproductive
41	289.5	20.6	358	22	AAAG26897	Hydrophobic domain
42	286.5	20.4	139	22	AAAB12131	Novel human diagno
43	267	19.0	607	21	AAAB12131	
44	231.5	16.5	264	22	ABG27003	
45	228	16.2	538	22	ABAB60174	Drosophila melanog

ALIGNMENTS

RESULT 1
AAV52386 standard; Protein: 268 AA.
XX AAV52386;
AC
XX
XX 09-FEB-2000 (first entry)
XX
XX Human transmembrane protein HP02000.
DE
XX
XX HP02000: transmembrane domain; liver; expression; homology;
KW organic cation transporter; drug excretion; antibody; assay reagent;
KW diagnostic marker; primer; probe; antisense; gene therapy;
KW agonist; antagonist; ligand; therapeutic.
XX
XX Homo sapiens.
OS
XX
XX WO9955862-A2.
PN
XX
XX 04-NOV-1999.
PD
XX
XX 27-APR-1999; 99WO-JP02226.
PF
XX
XX 28-APR-1998; 98JP-0119395.
PR
XX
XX (SAGA) SAGAMI CHEM RES CENT.
PA (PROT-) PROTEGENE INC.
XX
XX Kato S, Kimura T;
PI
XX
XX WPI: 2000-023358/02.
DR N-PSDB: AA238316, AA238317.
XX
XX Human proteins with transmembrane domains, involved in control of cell

PT proliferation and differentiation, useful for treating e.g. cancer or
 PT inflammation

PS Claim 1: Page 74-75; 11app; English.

XX
 XX
 CC This sequence represents the protein encoded by human cDNA clone
 CC HP02000 which is a 32 kD in size with two putative transmembrane
 CC domains. The cDNA was isolated from a human liver cell cDNA library, and
 CC from tissue localisation studies has been found to be expressed only in
 CC the liver. The protein has homology with the rat organic cation
 CC transporter (EMBL Accession No. Y09945) which is involved in drug
 CC excretion, and this sequence may have a similar function. The protein
 CC may be used to raise specific antibodies, as assay reagents, as
 CC ligands and binding proteins, for the isolation of cognate receptors,
 CC Nucleotides encoding the protein may be used as primers and probes or
 CC antisense molecules, and in gene therapy. Cells transformed with these
 CC nucleotides may be used to screen for agonists and antagonists which are
 CC potentially useful therapeutically.

XX Sequence 268 AA:

Query Match 100.0%; Score 1406; DB 21; Length 268;
 Best Local Similarity 100.0%; Pred. No. 1.2e-143; Indels 0; Gaps 0;
 Matches 268; Conservative 0; Mismatches 0;

QY 1 MAFELLISQVGLGRFOMLHVFLIPSLMLIPHLLENFAAIPGRCWVHMLDNTGS 60
 Db 1 mafeellisqvglgrfomlhvflipsmliphllefaaiphrcwvhlmdntgs 60
 QY 61 GNEGTLSLSEDAALRISIPDSNLRPEKCRREHPWOMQLHNGTHTSSEADTEPCVDGW 120
 Db 61 gnetgtslseadalrisipdsnlrpekcrrfhpwqllhngthtsseatdepcvdgw 120
 QY 121 VYDOSYFPSTIVTKMDLVCDYQSLKSVQFLLTGMVGGIIGHVSDDRIYVESARWLI 180
 Db 121 vydosyfpstivtkmdlvcdyqslksvqflltgmlyggihghvsdrivesarwli 180
 QY 181 TNKDEGKALKRKARVINGIKNAEETLNIEVVRSTMOEELDAOQRTTYVCDFRNSMRK 240
 Db 181 tnkdegalkrkarkarvngiknaeetlnievvrstmoeeldaoqrttyvcdfnrsmrk 240
 QY 241 RICILVLRKKIRSRHRNDCTYKTKF 268
 Db 241 ricilvlrkkirsrhrndctytktkf 268

RESULT 2

ABAB47276
 ID AAB47276 standard; Protein: 541 AA.

XX AAB47276;

DT 06-AUG-2001 (first entry)

XX hoAT5.

XX Human; organic anion transporter; hoAT; liver; kidney;
 KW membrane protein; transport; organic anion; splice variant.
 XX
 OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 282 /note= "PKC phosphorylation site"
 FT Modified-site 289 /note= "PKC phosphorylation site"
 FT Modified-site 345 /note= "PKC phosphorylation site"
 FT Modified-site 526 /note= "PKC phosphorylation site"
 FT /note= "PKC phosphorylation site"

PN WO200104283-A2.

PD 18-JAN-2001.

PF 12-JUL-2000; 2000WO-US18980.

PR 12-JUL-1999; 99US-0143771.

PA (META-) METABASIS THERAPEUTICS INC.

PI Sun W;

DR WPI: 2001-367057/38.

DR N-PSDB: AAC85824.

FT Nucleic acids encoding human organic anion transporter polypeptides,
 PT useful in gene therapy procedures -

PS Claim 2; Fig 6; 95pp; English.

XX
 XX The sequences given in AAB47271-76 represent human organic anion
 CC transporter (hoAT) polypeptides. hoAT polypeptides are preferentially
 CC expressed in the liver and kidneys of humans. OA's are membrane
 CC proteins that facilitate the transport of organic anions across the
 CC cell membrane. The mechanism of transport is thought to be a secondary
 CC or tertiary active transport involving exchange of another organic anion.
 CC hoAT2A and hoAT2B are thought to be splice variants as they are
 CC identical except at the C-terminal end. hoAT proteins and the DNA
 CC encoding them, may be used in the prevention, treatment and diagnosis
 CC of diseases associated with inappropriate hoAT expression.

SQ Sequence 541 AA:

Query Match 88.1%; Score 1238; DB 22; Length 541;
 Best Local Similarity 69.7%; Pred. No. 5.2e-125;
 Matches 249; Conservative 0; Mismatches 0; Indels 108; Gaps 1;

QY 1 MAFELLISQVGLGRFOMLHVFLIPSLMLIPHLLENFAAIPGRCWVHMLDNTGS 60
 Db 1 mafeellisqvglgrfomlhvflipsmliphllefaaiphrcwvhlmdntgs 60
 QY 61 GNETGTSLSLSEDAALRISIPDSNLRPEKCRREHPWOMQLHNGTHTSSEADTEPCVDGW 120
 Db 61 gnetgtslseadalrisipdsnlrpekcrrfhpwqllhngthtsseatdepcvdgw 120
 QY 121 VYDOSYFPSTIVTKMDLVCDYQSLKSVQFLLTGMVGGIIGHVSDDRIYVESARWLI 168
 Db 121 vydosyfpstivtkmdlvcdyqslksvqflltgmlyggihghvsdrivesarwli 168
 QY 169 -----
 Db 169 -----
 QY 181 IqlaIdcaafptlfvycvllrflagfssmltismnsipitewlrpnskalvvlssga 240
 Db 169 -----
 QY 241 lsgglllgglayvfrdwqclhvvaavpflglllgvrlvesarwllitknkdegkalkr 192
 Db 193 kvartngiknaeetlnievvrstmoeeldaoqrttyvcdfnrsmrkrcicilvflr 300
 QY 301 kvartngiknaeetlnievvrstmoeeldaoqrttyvcdfnrsmrkrcicilvflr 357
 Db 301 kvartngiknaeetlnievvrstmoeeldaoqrttyvcdfnrsmrkrcicilvflr 357

RESULT 3

ABAB11732
 ID ABB11732 standard; peptide; 391 AA.

XX ABB11732;

DT 11-JAN-2002 (first entry)

XX Human secreted protein homologue, SEQ ID NO:2102.
 DE
 XX

Human: cytokine; cell proliferation; cell differentiation; growth factor;
 haematopoiesis regulation; tissue growth; immunomodulator; activin;
 inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 chronic inflammatory condition; proliferative retinopathy;
 atherosclerosis; coronary heart disease; arterial ischaemia;
 bone disorder; osteoporosis; vascular growth disorder;
 tissue regeneration; wound healing; infection; immune disorder;
 cell culture; drug screening; gene therapy; antiinflammatory;
 antistatic; antiatheritic; haemostatic; antiarteriosclerotic;
 cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 antifungal; vunerary; antitumor.

XX Homo sapiens.
 XX WO200157188-A2.
 XX 09-AUG-2001.
 XX 05-FEB-2001; 2001WO-US03800.
 XX 03-FEB-2000; 2000US-0496914.
 XX 27-APR-2000; 2000US-0560875.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT;
 XX WPI: 2001-457740/49.
 XX N-PSDB: ABA08976.
 XX Human proteins and DNA encoding sequences useful for preventing,
 XX treating or ameliorating a medical condition in a mammalian subject
 XX e.g. arthritis and cancer -
 XX
 XX Claim 20; Page 235-236; 1963pp; English.

Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 invention also relates to vectors and recombinant host cells comprising a
 nucleotide of the invention, methods of producing the novel polypeptides,
 antibodies against the polypeptides, methods of detecting the nucleotides
 or polypeptides in a sample, and methods of identifying compounds which
 bind to polypeptides of the invention. Although novel, many of the
 polypeptides of the invention have homology to known proteins, thereby
 giving an insight into their probable biological activities, and hence
 potential therapeutic applications. The polypeptides of the invention may
 have various activities, including cytokine, cell proliferation or cell
 differentiation activities; stem cell growth factor activity;
 haematopoiesis regulatory activity; tissue growth activity;
 immunomodulatory activity; activin- or inhibin-related activities;
 chemotactic or chemokinetic activities; haemostatic, thrombotic or
 thrombolytic activities; receptor or ligand activities; or may be
 involved in oncogenesis, cancer cell proliferation or metastasis.
 Depending on their biological activities, polypeptides and nucleotides of
 the invention are useful for preventing, treating or ameliorating medical
 conditions, e.g., by protein or gene therapy. Such conditions include
 cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 proliferative retinopathy, atherosclerosis, coronary heart disease,
 arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 vascular growth. Polypeptides involved with tissue regeneration and
 repair (or nucleic acids encoding them) may be used to promote wound
 healing (e.g., of burns, incisions and ulcers), while those with
 immunomodulatory activities may be used in the treatment of viral,
 bacterial and fungal infections in addition to immune disorders.
 Polypeptides with growth factor activity may be used in cell cultures to
 promote cell growth. For example, such polypeptides may be used to
 manipulate stem cells in culture to give rise to neuroepithelial cells
 that can be used to augment or replace cells damaged by illness,
 autoimmune disease or accidental damage. The polypeptides and nucleotides
 may also be used in the diagnosis of the above conditions, and in drug

screening techniques. The present sequence represents a novel human
 polypeptide of the invention.

Sequence 391 AA;

Query Match 88.0%; Score 1237; DB 22; Length 391;
 Best Local Similarity 69.6%; Pred. No. 4,1e-125;
 Matches 250; Conservative 0; Mismatches 1; Indels 108; Gaps 1;

1 MAFPELLSQVGLGRFOWLHLVFIPLSLMLIPHLLENFAAATPGHRCWVHMDNNTGS 60
 27 mafeellsqvglggrfgmhlhvfplslmlliphllefaaapghrcwvhmdnntgs 86
 61 GNETGIISEDALRLTSLPLDSNLRPEKCRVYHPQWOLHNGTHTSHSEADTEPCVDGW 120
 87 gnetgllsedallrltslpldsnlrpekcrvthpqqwllhngtthsthsdeadtepcvdgw 146
 121 VYDOSYFPESTIVTKMDLVCDYQSLKSVQFLLLTGMLVGIGCHVSD----- 168
 147 vydgyfptstlvtkwldvcdyqslksvqfllltgmlvgigghvsdrfgrrllrwgl 206
 169 ----- 168
 207 lqalttcaafaptfpyovrlflagfsmllsnnslptewlrpnskalvvlssga 266
 169 -----RMVESA RMLITIRKLDGKALR 192
 267 lnlgglilgglayfirdwqlhvasvpffvflrlstvwesarvlltnkldgalkar 326
 193 KYAPNGIKNAEETINIEVVRSTMOEELDAQTKTVCDLFNPNRKRKICILVFLRK 251
 327 kyatnglknasetlnievvrstmqeeldaqtktvcdlfnpnmrkrkicllvflrk 385

RESULT 4
 AAM24188
 ID AAM24188 standard; Protein; 213 AA.
 XX
 AC AAM24188;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE Human EST encoded protein SEQ ID NO: 1713.
 XX
 KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
 KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
 KW diagnostics; forensic test; gene mapping; genetic disorder;
 KW biodiversity; gene therapy; nutrition.
 XX
 OS Homo sapiens.
 XX
 PN WO200154477-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 25-JAN-2001; 2001WO-US02687.
 XX
 PR 25-JAN-2000; 2000US-0491404.
 PR 17-JUL-2000; 2000US-0617746.
 PR 03-AUG-2000; 2000US-0631451.
 PR 15-SEP-2000; 2000US-0663870.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
 PI Cao Y, Drmanac RA, Zhang J, Werhman T;
 XX
 DR WPI: 2001-476164/51.
 XX N-PSDB: AAH98847.
 XX
 PT Isolated polypeptide for treatment of diseases, diagnostics, raising
 PT antibodies and research use -

XX Claim 20; Page 1141; 1275bp; English.

XX

CC The present invention provides the protein and coding sequences of novel
CC proteins from a variety of organisms, including human, dog, cat, horse,
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC from the organism of interest. They can be used in diagnostics,
CC forensics, gene mapping, identification of mutations, to assess
CC biodiversity and for nutritional purposes. The present sequence is a
CC protein of the invention.

50 Sequence 213 AA;

Query Match	70.6%	Score 993	DB 22	Length 213
Best Local Similarity	100.0%	Pred. No. 4,46	99	
Matches 187	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY	18	MLHVFELPSMLLIPHLLENFPAALIDGHHCWVHMLDNNTGSGNETGISSEDAALRISI	77
Db	1	mlhvfllpsmliphllelntfaaapghrcwvbmldnntgsgnetgisledallrlisi	60
QY	78	PLDSNLREPKCRREYHPOMOLHLNGTHTSHSEADTEPCVDGWVYDQSYRPSSTIVLTWDL	137
Db	61	pldsnlrepekerrtyhbpqwlrlhngtlhshseadtepcvdgwvydqsyrpsstvtkwdl	120
QY	138	VCDYQSLKSVVQFLLLTGMVLGGIIGGHSVDRLVLESARMLITYNKIDEGKALRKVART	197
Db	121	vcdyqslksvqvqlllfgmlvygslggghvdsrvtvesarwllitnkldegikalkrvart	180
QY	198	NGRIINAE 204	
Db	181	ngriinae 187	

RESULT	5
AAE04897	
ID	AAE04897 standard; Protein; 553 AA.

AC AAE04897;

DT 10-SEP-2001 (first entry)

Human transporter and ion channel-10 (TRICH-10) protein.

KW Human; transporter and ion channel-10; TRICH-10; cystic fibrosis; mood;
KW gene therapy; amyotrophic lateral sclerosis; amnesia; muscular dystrophy;
KW hypertension; angina; neurological disorder; asthma; bipolar disorder;
KW dementia; depression; Alzheimer's disease; epilepsy; vaccine; arhythmia
KW Pick's disease; Ischemic cerebrovascular disease; AIDS; anxiety; stroke
KW Huntington's disease; Parkinson's disease; cerebral neoplasm; allergy;
KW demyelinating disease; mental disorder; Schizophrenia; polymyositis;
KW muscle disorder; cardiomyopathy; cataract; myocarditis; Grave's disease;
KW dermatomyositis; diabetes mellitus; immunological disorder; psoriasis;
KW Rheumatoid arthritis; Sjogren's syndrome; systemic lupus erythematousus;
KW sickle cell anaemia; Wilson's disease; infertility; Cushing's disease;
KW scleroderma; pulmonary artery stenosis; neutropenia; Addison's disease;
KW malabsorption syndrome; hypercholesterolaemia; cancer.

Homo sapiens.

Key	Location/Qualifiers
EH	
EE	

Domain	204. .222
--------	-----------

```
/label= Transmembrane_domain
```

Domain	470..493
--------	----------

```

/label= Transmembrane_domain

```

Domain	500..519
ET	

```

/label= Transmembrane_domain

```

WO200146258-A2

PD 28-JUN-2001.

XX
PF 22-DEC-2000; 2000WO-US35095.
VV

PR 23-DEC-1999; 99US-0172000
PR 14-JAN-2000; 2000US-0176083
PR 21-JAN-2000; 2000US-0177332
PR 28-JAN-2000; 2000US-0178572
PR 02-FEB-2000; 2000US-0179758
PR 10-FEB-2000; 2000US-0181625.

PA (INCY-) INCYTE GENOMICS INC.

PI Baughn MR, Butford N, Au-Young J, Lu DAM, Yang J, Reddy R, Lal P,
PI Hillman JL, Azimzai Y, Yue H, Nguyen DB, Yao MG, Gandhi AR,
PI Tang YT, Khan FA;

DR WPI; 2001-418042/44.
DR N-PSDB; AAD09561.

PT Novel human transporter and ion channel proteins useful for treating
PT and preventing transport, neurological, muscle and immunological
PT disorders -

PS Claim 1; Page 122-123; 160pp; English.

The present sequence is transporter and ion channel-10 (TRICH-10) protein. TRICH is used as vaccine. TRICH is useful for treating a disease or condition associated with decreased expression of functional TRICH, such as transport disorder including amyotrophic lateral sclerosis, cystic fibrosis, Becker's muscular dystrophy, Charcot-Marie Tooth disease, Duchenne muscular dystrophy, angina and hypertension, neurological disorders including Alzheimer's disease, amnesia, bipolar disorder, dementia, depression, epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Pick's disease, Huntington's disease and Parkinson's disease, demyelinating diseases, mental disorders including mood, anxiety, schizophrenia and seasonal affective disorder, muscle disorder including cardiomyopathy, myocarditis, polymyositis, dermatomyositis, arthylmitias and asthma and immunological disorders, including AIDS, adult respiratory distress syndrome (ARDS), allergies, anaemia, diabetes mellitus, rheumatoid arthritis, scleroderma, Sjogren's syndrome, systemic lupus erythematosus and other diseases including sickle cell anaemia, Wilson's disease, cataracts, infertility, pulmonary artery stenosis, Grave's disease, Cushing's disease, Addison's disease, glucose-galactose malabsorption syndrome, hypercholesterolaemia, cancers psoriasis and viral, bacterial, fungal, helminthic and protozoal infections. TRICH DNA is useful in gene therapy and in diagnostic purposes.

SQ Sequence 553 AA;

Query Match	56.98	Score 799.5	DB 22	Length 553
Best Local Similarity	44.7%	Pred. No. 1.7e-77		
Matches 160	Conservative 40	Mismatches 49	Indels 109	Gaps 2

1 MAEEELLISVGGIGRPFOMLHVFLLPSIMLLPHILENNAALPGHRGCVHMLDNNTGS 60
 ||||| | | ||| | | : | : |||| | ||||| | : ||| |
 1 mafqdlqhnagd|wrfql|qvfvis|favaty|hmlentat|phrcwv|hldndtvs 60

[illegible]

121 VDDSYFSTIVTKWDLVCDYQSLKSVQFLLTGMLVGIIIGHVSDR----- 169

[illegible]

121 vyarissstivewdlvcdsqsltsvakfvfmgmmvvgllgghlsdrfgrrfvllrwcyl 180

170 -----

I/O	-----	169
22	-----	

```
181 j'avai vatcaajant flijvcs] r f] s a i a a m s j i t o t i m j s c o r t b a f e z m m e i 4 j ----- 2 1 0
```

101 qvavvccaaaplllycsllrllsglaamslltntlmllaewathrfqamgltlmgcp 240

```

2Y 170 -----WLVESARWLIITNKIDEGIKALR 192

```


Db 241 sgiafmltaglafairdwhlqlvsvpyfvfltsawllesarwllinnkpeegjkeir 300
 QY 193 KVARINGIKNAEETINIEVVRSTMOEELDAO-TKTTVCDFRNPSMKRRICTIVFLR 249
 301 kaahrsgmknardltlletlksctmkleaaqkkspsicemlmpnckrislstr 358

RESULT 6
 AAB47275
 ID AAB47275 standard; Protein; 554 AA.

AC AAB47275;
 DT 06-AUG-2001 (first entry)

DE hOAT4.

Human; organic anion transporter; hOAT; liver; kidney;
 membrane protein; transport; organic anion; splice variant.

OS Homo sapiens.

Key Location/Qualifiers
 FT Modified-site 324 /note="PKC phosphorylation site"

MO200104283-A2.

18-JAN-2001.

12-JUL-2000; 2000MO-US18980.

12-JUL-1999; 99US-0143771.

(META-) METABASIS THERAPEUTICS INC.

PI Sun W;

WPI; 2001-367057/38.

N-PSDB; AAC85823.

Nucleic acids encoding human organic anion transporter polypeptides,
 useful in gene therapy procedures.

Claim 2; Fig 5; 95pp; English.

The sequences given in AAB47271-76 represent human organic anion
 transporter (hOAT) polypeptides. hOAT polypeptides are preferentially
 expressed in the liver and kidneys of humans. OAT's are membrane
 proteins that facilitate the transport of organic anions across the
 cell membrane. The mechanism of transport is thought to be a secondary
 or tertiary active transport involving exchange of another organic anion.
 hOAT2A and hOAT2B are thought to be splice variants as they are
 identical except at the C-terminal end. hOAT proteins and the DNA
 encoding them, may be used in the prevention, treatment and diagnosis
 of diseases associated with inappropriate hOAT expression.

Sequence 554 AA:

Query Match 54.6%; Score 768; DB 22; Length 554;
 Best Local Similarity 43.7%; Pred. No. 4,4e-74;
 Matches 157; Conservative 40; Mismatches 52; Indels 110; Gaps 3;

QY 1 MAREELLQGVGGIRFOMLHVFILPSMLLIPHLLENFAAIPGRCVWMLDNNTGS 60

Db 1 matgdllghgdltwriglqvtflsifavatyhfmntafipghrcwvhlldnlys 60

QY 61 GNETIGLSEDAALRISIPDLSNRPEKCRFVHPQOLHLNNTISTSPADPEPCVDGW 120

Db 61 dndtgalsgdallrtisipdlsnmrpekcrtfivpqwllhnglftpsdaamepcvdgw 120

QY 121 VYDSYFPPTI-VTKMVLVDYQSLKSVQFLLLTGMVLGGIIGHVSDR----- 169

Db 121 yldrsfslsdglkwdlvcdsqslsvakfvmagmmllgdlghlsdrgrfsvlrc 180

QY 170 ----- 169

Db 181 ylqvalvgcaalaptfllycsvrfllsgiaamsflntlmllaewathrfqamgtlqmc 240

QY 170 ----- 191

Db 241 psgiafmltaglafairdwhlqlvsvpyfvfltsawllesarwllinnkpeegjkeir 300

QY 192 KVARINGIKNAEETINIEVVRSTMOEELDAO-TKTTVCDFRNPSMKRRICTIVFLR 249

301 rkaahrsgmknardltlletlksctmkleaaqkkspsicemlmpnckrislstr 359

RESULT 7
 AAM40059
 ID AAM40059 standard; Protein; 341 AA.

AC AAM40059;

DT 22-OCT-2001 (first entry)

Human polypeptide SEQ ID NO 3204.

Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 peripheral nervous system neuropathy; central nervous system; CNS;
 Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 leukaemia.

OS Homo sapiens.

PI WO200153312-A1.

26-JUL-2001.

26-DEC-2000; 2000MO-US34263.

21-JAN-2000; 2000US-0488725.

PR 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

(HYSE-) HYSEQ INC.

Tang Y, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 Zhao QA, Zhou P, Goodrich R, Drmanac R;

WPI; 2001-442253/47.

N-PSDB; AAI59215.

Novel nucleic acids and polypeptides, useful for treating disorders
 such as central nervous system injuries.

Example 5; SEQ ID NO 3204; 10078pp; English.

The invention relates to human nucleic acids (AA157798-AA161369) and
 the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 immunosuppressant and cytostatic activity. The polynucleotides are useful
 in gene therapy. A composition containing a polypeptide or polynucleotide
 of the invention may be used to treat diseases of the peripheral nervous
 system, such as peripheral nervous injuries, peripheral neuropathy and
 localized neuropathies and central nervous system diseases, such as
 Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.

XX Sequence 341 AA:

Query Match 54.4%: Score 765; DB 22; Length 341;
 Best Local Similarity 45.5%: Pred. No. 4.5e-74;
 Matches 152; Conservative 33; Mismatches 41; Indels 108; Gaps 1;

OY 1 MAFPELLSQVGLGRFQMLHVFILPSLMLLPHILENFAAIPGHCWVHMDNNTGS 60
 Db 1 mafgdlighagdlwrfqllqvfisifavatyihfmentafipghrcwvhlidndtvs 60
 OY 61 GNETGILSEDLALRISIPIDSNLRPEKCRFVHPQWOLHNTIHSTSEADTEPCVDGW 120
 Db 61 dndtgalsgdallrlstipidsnmrpekrffvhpqwlhlnghfptsdadmepcvdgw 120
 OY 121 VYDOSYFSPSTIYTKMDLVCDYQSLKSVQFLLTGMVGGIIGGHVSDR 169
 Db 121 vydrisfsaisvewdlvcdsgsltsvakfvfmgamvvgllgshdsdrrfrrfwlrcwy 180
 OY 170 ----- 169
 Db 181 lqvalvgtcaalapfllyclrlfsgiaamslntntlmiaewathrfqangiltmcp 240
 OY 170 ----- 169
 Db 241 sglaftmclaglaafalrdwhllqlvsvpyfvflfletswllsesarvllinnkpeegikrlr 192
 OY 193 KVARINGIKNAEETLINIEVVRSTMOEELDAQTK 226
 Db 301 kahtsgmknardctltleilkstmkleaaqkk 334

RESULT 8

AAE19406
 ID AAE19406 standard; Protein; 339 AA.

AC AAE19406;

DT 06-MAR-2001 (first entry)

DE Amino acid sequence of a human secreted protein.

KW Secreted protein; platelet disorder; stem cell disorder; osteoporosis;
 KW osteoarthritis; burn; incision; ulcer; periodontal disease; neuropathy;
 KW nervous system disease; bone growth; cosmetic plastic surgery;
 KW gut protection; gut regeneration; fibrosis; cancer;
 KW bone marrow transplantation.

OS Homo sapiens.

PN WO200061755-A2.

PD 19-OCT-2000.

PE 10-APR-2000; 2000MO-US09555.

PR 09-APR-1999; 99US-0128574.

PR 20-AUG-1999; 99US-0150054.

PA (CHIR) CHIRON CORP.

PI Garcia PD;

DR WPI; 2000-665133/64.

DR N-PSDB; AAC61892.

XX Novel secreted human proteins useful for stimulating blood cell
 PT generation in patients receiving cancer chemotherapy, treating bone
 PT marrow transplantation patients and for healing fractured bones
 XX Claim 14; Page 72; 74pp; English.

AAE19393-B19407 represent secreted human proteins. The secreted proteins
 CC are useful in assays to determine their biological activities. The
 CC proteins can also be used as biomarkers to identify tissues or cell
 CC types which express the proteins. The polynucleotide molecules or cell
 CC used as biomarkers for tissues or chromosomes and to elicit immune
 CC responses. The proteins and antibodies are useful in diagnosis and
 CC treatment of diseases associated with altered expression of these
 CC proteins. The proteins are also useful for prevention or treatment of
 CC platelet disorders, stem cell disorders, osteoporosis or osteoarthritis,
 CC burns, incisions, ulcers, periodontal diseases, central and peripheral
 CC nervous system diseases and neuropathies, for healing fractured bones
 CC and to induce cartilage and/or bone growth in cosmetic plastic surgery.
 CC The proteins are also useful for gut protection or regeneration, for the
 CC treatment of lung or liver fibrosis, for stimulating blood cell
 CC generation in patients receiving cancer chemotherapy and for treatment
 CC of bone marrow transplantation patients.

XX Sequence 339 AA:

Query Match 54.0%: Score 759; DB 21; Length 339;
 Best Local Similarity 44.9%: Pred. No. 2e-73;
 Matches 151; Conservative 33; Mismatches 44; Indels 108; Gaps 1;

OY 1 MAFPELLSQVGLGRFQMLHVFILPSLMLLPHILENFAAIPGHCWVHMDNNTGS 60
 Db 1 mafgdlighagdlwrfqllqvfisifavatyihfmentafipghrcwvhlidndtvs 60
 OY 61 GNETGILSEDLALRISIPIDSNLRPEKCRFVHPQWOLHNTIHSTSEADTEPCVDGW 120
 Db 61 dndtgalsgdallrlstipidsnmrpekrffvhpqwlhlnghfptsdadmepcvdgw 120
 OY 121 VYDOSYFSPSTIYTKMDLVCDYQSLKSVQFLLTGMVGGIIGGHVSDR 169
 Db 121 vydrisfsaisvewdlvcdsgsltsvakfvfmgamvvgllgshdsdrrfrrfwlrcwy 180
 OY 170 ----- 169
 Db 181 lqvalvgtcaalapfllyclrlfsgiaamslntntlmiaewathrfqangiltmcp 240
 OY 170 ----- 169
 Db 241 sglaftmclaglaafalrdwhllqlvsvpyfvflfletswllsesarvllinnkpeegikrlr 192
 OY 193 KVARINGIKNAEETLINIEVVRSTMOEELDAQTK 226
 Db 301 kahtsgmknardctltleilkstmkleaaqkk 336

RESULT 9

AAE10332
 ID AAE10332 standard; Protein; 553 AA.

AC AAE10332;

DT 10-DEC-2001 (first entry)

DE Human transporter and ion channel-9 (TRICH-9) protein.

KW Human; transporter and ion channel; TRICH-9; therapy; akinesia; cardiac;
 KW neurological disorder; immune disorder; allergy; nocturnal dementia;
 KW AIDS; Acquired Immune Deficiency Syndrome; amyotrophic lateral sclerosis;
 KW cystic fibrosis; Alzheimer's disease; depression; epilepsy; scleroderma;
 KW cardiomyopathy; muscular dystrophy; Crohn's disease; multiple sclerosis;
 KW rheumatoid arthritis; cancer; anaemia; asthma; trauma; haemodialysis;
 KW gastritis; inflammation.

XX Homo sapiens. Location/Qualifiers
 OS 106..530
 Key /label= Sugar transporter
 FH 204..222
 Domain /label= Sugar_transport-protein
 FT 357..383
 FT /label= Sugar_transport-protein
 Domain
 WT000162923-A2.
 PN 30-AUG-2001.
 PD 23-FEB-2001; 2001WO-US05942.
 PE 25-FEB-2000; 2000US-0184866.
 PR 02-MAR-2000; 2000US-0187947.
 PR 09-MAR-2000; 2000US-0188333.
 PR 17-MAR-2000; 2000US-0190230.
 PR 24-MAR-2000; 2000US-0192077.
 PR 30-MAR-2000; 2000US-0193500.
 PA (INCYTE) INCYTE GENOMICS INC.
 PI Yue H, Tang YF, Lal P, Policky JT, Nguyen DB, Au-Young J, Yao MG;
 PI Khan FA, Walla NK, Gandhi AR, Tribouley CM, Patterson C;
 PI Thornton M, Greene BD, Hernandez R, Borowsky ML, Sanjanwala MS;
 DR N-PSDB; AAD17476.
 XX WPI: 2001-582050/65.
 XX Thirteen human transporters and ion channels (referred to as TRICH-1 to
 PT TRICH-13), useful in the diagnosis, treatment and prevention of
 PT transport (e.g. akinesia), neurological, muscle or immunological
 PT disorders (e.g. allergies) -
 PS Claim 1; Page 114-115; 131pp; English.
 CC The present sequence is human transporters and ion channels
 CC (TRICH-9) protein. The TRICH DNA, protein and their agonist and
 CC antagonists are useful in the diagnosis, treatment and prevention of
 CC transport disorders (akinesia, amyotrophic lateral sclerosis, cystic
 CC fibrosis), neurological (Alzheimer's disease, dementia, depression) or
 CC epilepsy), muscle (e.g. cardiomyopathy, Becker's muscular dystrophy) or
 CC immunological disorders (e.g. allergies, acquired immunodeficiency
 CC syndrome (AIDS), Crohn's disease, rheumatoid arthritis, scleroderma,
 CC multiple sclerosis), viral bacterial, parasitic, protozoal and
 CC helminthic disorders, cancer, anaemia, asthma, trauma, haemodialysis,
 CC gastritis and inflammation.
 XX Sequence 553 AA;
 SQ 42.7%; Score 601; DB 22; Length 553;
 Query Match 37.8%; Pred. No. 5,3e-56; Indels 108; Gaps 1;
 Match 133; Conservative 30; Mismatches 81;
 QY 1 MAFEELISQVGGGRFQMLHYFLIPSLMLPILLENFAALIPGRVHMIDNNTGS 60
 DB 1 mafeellidvggldvgrfyqfqlmalmwlmwctqsmlefaaapsrncwappildnstaq 60
 QY 61 GNGTGLSDALIRISIPDSNMRKCKRREYHPQWOLHNGTTHSTSEADTEPCVGDW 120
 DB 61 asllgelspeallalsipgpnpqrqgrtrfqpqwdldpnatstseadtepcvqdw 120
 QY 121 VYDOSYFPSTIVKKNLNDYOSLKSVOFLILTGMLIGGIVSRR 169
 DB 121 vdsitstivakmnlvcshalkpmagsiyagllvgaaacpssadrfgrtrlyltsy 180
 QY 121 vdsitstivakmnlvcshalkpmagsiyagllvgaaacpssadrfgrtrlyltsy 180
 QY 170 ----- 169

DB 181 lqamvmtaaafapafpyclffllafavaymmnltgllmewtaarlpvmtlnslg 240
 QY 170 -----WVESARWLITNKIDEGIKAR 192
 DB 241 fefghlaavaygrdcllqlvsvpfllcflyswlaesarwlltgrldwqlelw 300
 QY 193 KYARNGIKNAEETLINEVRSSTMOELDAACOTTYCDEFRNPSMKRRICI 244
 DB 301 rvaalngkxavqdltplevllsaameismqgppaslgllimpqifrccl 352
 RESULT 10
 ID AAM41845 standard; protein; 369 AA.
 AC AAM41845;
 XX 22-OCT-2001 (first entry)
 DE Human polypeptide SEQ ID NO 6776.
 XX Human: nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 OS Homo sapiens.
 XX W0200153312-A1.
 PN 26-JUL-2001.
 PD 26-DEC-2000; 2000WO-US34263.
 PE 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 PA (HYSEQ) HYSEQ INC.
 PI Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 DR WPI: 2001-442253/47.
 DR N-PSDB; AAI61001.
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 XX such as central nervous system injuries -
 PS Example 2; SEQ ID NO 6776; 10078pp; English.
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM438642-AA42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Actin/Inhibin activity, Chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 SQ Sequence 369 AA;

Query Match
 Best Local Similarity 42.3%; Score 594.5; DB 22; Length 369;
 Matches 123; Conservative 30; Mismatches 32; Indels 109; Gaps 2;

QY 65 GILSEDLRLRISIPDSNLRPEKCRREVHOMQLHNGTISTSEADTEPCVDGWYDQ 124
 Db 9 galsqdaillrlslpidsnmpekcrrivhpqwlhngfprtsdadmepcvdgywdt 68
 QY 125 SYPSRTVTKWDVLCYQSKSVYQPLLTGMVGGIIGCHVSDR----- 169
 Db 69 lsfstlvekwldvcdsqsltsvaktfvnmvgllghlsdtrgrfvrlwcylyva 128
 QY 170 ----- 169
 Db 129 lvtcaalaprfllycsrlflsgiaamslltntlllaewathrfqamgiltgmcpsgla 188
 QY 170 ----- 169
 Db 189 fmltlaqlafairdwhllglvsvpyfvflfclswlllesarwlhmkpgeiketrkaah 196
 QY 197 TNGIKNAEETINIEVNSTMOEELDAQOTKT-VCDFLRNPMKRCILVFLR 249
 Db 249 rsgmknardltleilkstmkleaaqkkpflgerlmpnckrtslipfk 302

RESULT 11

ID AAE06612 standard; Protein; 550 AA.
 AC AAE06612;

25-SEP-2001 (first entry)

Human protein having hydrophobic domain, HP03882.

KW Human; hydrophobic domain; gene therapy; nutritional supplement;
 KW multiple sclerosis; immunomodulatory; autoimmune disorder; antimicrobial;
 KW haematopoiesis; tissue growth activity; insulin-dependent diabetes;
 KW Huntington's disease; Alzheimer's disease; Parkinson's disease; cytostatic;
 KW haemostatic; thrombolytic; tumour growth inhibitor; anabolic;
 KW contractile; antinflammatory; antinflammatory;
 KW Homo sapiens.

OS
 XX WO200149728-A2.

PD 12-JUL-2001.

PF 28-DEC-2000; 2000WO-JP09359.

PR 06-JAN-2000; 2000JP-0000585.

PR 06-JAN-2000; 2000JP-0000588.

PR 03-FEB-2000; 2000JP-0002682.

PR 03-MAR-2000; 2000JP-0058367.

PA (PROT-) PROTEGENE INC.
 PA (SAGA) SAGAMI CHEM RES CENT.

PI Kato S, Kimura T;

WPI: 2001-418355/44.

DR N-PSDB; AADI2607.

PT Human proteins with hydrophobic domains and the nucleic acids encoding
 them, useful for preventing diagnosing and treating e.g. cancer.

PT Alzheimer's and inflammation -

PS Claim 1; Page 137-138; 563pp; English.

CC The present sequence is human protein with hydrophobic domain,
 CC HP03882. The polynucleotide and polypeptide of the invention
 CC may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate polypeptide expression. The polynucleotides
 CC may be used to produce the polypeptide, by inserting the nucleic acids
 CC into a host cell and culturing the cell to express the protein. The
 CC polynucleotides and its complementary sequences may also be used as DNA
 CC probes in diagnostic assays and also used in gene therapy. The
 CC and in assays to identify modulators of polypeptide expression and
 CC activity. The polypeptides and nucleic acids may be used as nutritional
 CC supplements, to modulate cytokine and cell proliferation activity, to
 CC modulate immune stimulation or suppression (e.g. for the treatment of
 CC microbial infections and autoimmune disorders such as multiple sclerosis,
 CC rheumatoid arthritis and insulin-dependent diabetes), to modulate
 CC treatment of Parkinson's disease, Huntington's disease and Alzheimer's
 CC disease), to modulate cytokine and inhibit activity (e.g. for controlling
 CC modulate haemostatic and thrombolytic activity, to modulate receptor
 CC ligand activity, to modulate inflammation and to inhibit tumour growth.
 CC Sequence 550 AA.

Query Match
 Best Local Similarity 42.1%; Score 592; DB 22; Length 550;
 Matches 143; Conservative 28; Mismatches 69; Indels 116; Gaps 7;

QY 1 MAFELISQVGLGFRFQWL-LVFLPSLMLIP-HILLENFAAATGHCWYMDNNT 58
 Db 1 mafeklllegagvgylfqltlytlpdlm--lpsgmllensaadlphhrcwltmdn-- 56
 QY 59 GSGNERTGILSEDLRLRISIPDSNLRPEKCRREVHOMQLHNGTISTSEADTEPCVD 118
 Db 57 gsnvstn-mrpkallstlshppnpgbqhgrtrfrpqwqldpnaatatswseadtepcvd 115
 QY 119 GWYDOSYPSRTVTKWDVLCYQSKSVYQPLLTGMVGGIIGCHVSDR----- 169
 Db 116 gwyydsvstfslvakwdlvcsgqlkplsgslfmgslfvgllsyrfgrkpmjsw 175
 QY 170 ----- 169
 Db 176 colqlavagstslfaptfviyglrfvaafmagflfsltlmwewttsrravltmlyvg 235
 QY 170 ----- 169
 Db 236 calsaggaalglafalrdwtlqlaasvpffalslswllpesarwlhmkpgeiketrkaah 190
 QY 191 LKRVARTGINKNAEETINIEVNSTMOEELDAQOTKTTCVDFLRNPMKRCILV 246
 Db 296 lrvkaringhkea-knltlevlmsvkeavasakprrsvldlfcvplwtrscaml 350

RESULT 12

ID AAB69091 standard; Protein; 550 AA.
 AC AAB69091;

24-APR-2001 (first entry)

Human organic anion transporter OAT4 protein sequence SEQ ID NO:2.

KW Human; organic anion transporter; OAT4; nephrotoxic; kidney disease;
 KW abnormal foetal growth.
 KW Homo sapiens.

QY 170 ----- 169
 Db 176 cclqlavagstlftaptfviyqglrfvaafmaglflslmlmewttsrravmtlvvg 235
 QY 170 -----WVESRRRLITNKDEGIKA 190
 Db 236 calsgagaalagglafalrdrtrtqlaasvptfaistlswlpsarvllkypdgaige 295
 QY 191 LRKARINGINNAEETLNIEVVRSTMOEELDAQTCTTCDLFNPSMRKRICLIVFLR 249
 Db 296 lrkvaringhkeetecvylkvlmsvaevasakeprsvldlfcvpylwrscamlvyk 354

RESULT 14

AAE06571
 ID AAE06571 standard; Protein: 578 AA.

AC AAE06571;

XX 25-SEP-2001 (first entry)

DE Human protein having hydrophobic domain, HP03613.

KW Human: hydrophobic domain; gene therapy; nutritional supplement;
 KW cell proliferation; immunomodulatory; autoimmune disorder; antimicrobial;
 KW multiple sclerosis; rheumatoid arthritis; insulin-dependent diabetes;
 KW haematopoiesis; tissue growth activity; Parkinson's disease; cytostatic;
 KW Huntington's disease; Alzheimer's disease; chemotactic; chemokinetic;
 KW contractile; thrombolytic; tumour growth inhibitor; anabolic;
 KW antiinfective; antiinfectivity; antiinflammatory.

OS Homo sapiens.

XX WO200149728-A2.

XX 12-JUL-2001.

XX 28-DEC-2000; 2000WO-JP09359.

XX 06-JAN-2000; 2000JP-0000585.

XX 06-JAN-2000; 2000JP-0000588.

XX 11-JAN-2000; 2000JP-0002299.

XX 03-FEB-2000; 2000JP-0026862.

XX 03-MAR-2000; 2000JP-0058367.

XX (PROT-) PROTEGENE INC.

XX (SAGA) SAGAMI CHEM RES CENT.

XX Kato S, Kimura T;

XX WPI; 2001-418355/44.

XX N-PSDB; AAD12566.

XX Human proteins with hydrophobic domains and the nucleic acids encoding

XX them, useful for preventing diagnosing and treating e.g. cancer,

XX Alzheimer's and inflammation -

XX Claim 1; Page 59-60; 563pp; English.

XX The present sequence is human protein with hydrophobic domain, HP03613.
 CC The polynucleotide and polypeptide of the invention may be used in the
 CC prevention, diagnosis and treatment of diseases associated with
 CC inappropriate polypeptide expression. The polynucleotides may be used
 CC to produce the polypeptide, by inserting the nucleic acids into a host
 CC cell and culturing the cell to express the protein. The polynucleotides
 CC and its complementary sequences may also be used as DNA probes in
 CC diagnostic assays and also used in gene therapy. The polypeptides may
 CC also be used as antigens in the production of antibodies and in assays
 CC to identify modulators of polypeptide expression and activity. The
 CC polypeptides and nucleic acids may be used as nutritional supplements,
 CC to modulate cytokine and cell proliferation activity, to modulate
 CC immune stimulation or suppression (e.g. for the treatment of microbial

CC infections and autoimmune disorders such as multiple sclerosis,
 CC rheumatoid arthritis and insulin-dependent diabetes), to modulate
 CC haematopoiesis, to modulate tissue growth activity (e.g. for the
 CC treatment of Parkinson's disease, Huntington's disease and Alzheimer's
 CC disease), to modulate activin and inhibin activity (e.g. for controlling
 CC fertility), to modulate chemotactic and chemokinetic activity, to
 CC modulate haemostatic and thrombolytic activity, to modulate receptor
 CC ligand activity, to modulate inflammation and to inhibit tumour growth.
 XX Sequence 578 AA;

Query Match 41.9%; Score 588.5; DB 22; Length 578;
 Best Local Similarity 35.3%; Pred. NO. 1.3e-54;
 Matches 133; Conservative 30; Mismatches 81; Indels 133; Gaps 1;

QY 1 MAFELLSSQVGGIGRPFOMLHVFLPSLMLDIPHLLENFAAIPGRVWHLMDNNTGS 60
 Db 1 mafellldlvvgllgrfvgllqumalmslmwlcqsmlemtsaavshrcwaplldnstaq 60
 QY 61 GNETGILSEDALIRISIPLDNSNLREPKRFRVHPQWOLHNLNGTISTSEADTEPCVDGM 120
 Db 61 asllglspeallaisippgpnqphqcrtrfipgqqlldpnatatsweadtepcvdgm 120
 QY 121 VPDOSVPSPIVKKMLVCDYOSLKSVOFLTLTGMVLGGIIGHSR----- 169
 Db 121 vpdosvpspiivkkmlvcdyoslksvoflltgmvlggiighsr----- 169
 QY 170 ----- 180
 Db 170 ----- 169
 Db 181 lqmvaymgtaaafapafpyclfrflafavagvmmntgltrsltwfhagglhgsraep 240
 QY 170 ----- 169
 Db 241 lglavmewtaaraaplvmctlnslgfsfghltaavaygyrdwllqlvsvpfflclly 300
 QY 170 --WVESRRRLITNKDEGLKALKVARTNGINNAEETLNIEVVRSTMOEELDAQTKT 227
 Db 301 swlaesarvlltrgrldwglgelvrvaalngkavagvcltpevlisameelsmgqppa 360
 QY 228 TVCDLFNPSMRKRIC 244
 Db 361 slgtllimpglfrfctl 377

RESULT 15

AAU32379

XX AAU32379 standard; Protein: 645 AA.

XX AAU32379;

XX 18-DEC-2001 (first entry)

XX Novel human secreted protein #2870.

XX Human: vaccination; gene therapy; nutritional supplement;

XX stem cell proliferation; haematopoiesis; nerve tissue regeneration;

XX immune suppression; immune stimulation; anti-inflammatory; leukaemia.

XX Homo sapiens.

XX WO200179449-A2.

XX 25-OCT-2001.

XX 16-APR-2001; 2001WO-US08656.

XX 18-APR-2000; 2000US-0552929.

XX 26-JAN-2001; 2001US-0770160.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX DR WPI: 2001-611725/70.

XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic
XX PR vaccination, testing and therapy -

PS Claim 20: Page 597; 765pp; English.

CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.

XX SQ Sequence 645 AA:

Query Match 33.6%; Score 472; DB 22: Length 645;
Best Local Similarity 28.8%; Pred. No. 6.3e-42;
Matches 126; Conservative 31; Mismatches 63; Indels 218; Gaps 6;

QY 25 LPSMLLIP-HILENFAAIPGHRWVHMLDNTSGNETGILSEDAALRISIPDSNL 83
DB 1 lpcdm--lpsqmllefaaiaiphrwcwchmdn--gsavstn-mtpkalltisipgpnuq 55
QY 84 RPEKCRFVHPQWQLLNHNGTHTSTSEADTEPCVYGWYIDOSTFPTITVKMDLYCDYQS 143
DB 56 gphgrrfrfqpqwlldpnatatsweadtepcvdgwydrsvftslvakwdlycsssqg 115
QY 144 LKSVQFLLTNGMLVGGIIGHVSDR----- 169
DB 116 lkpisqslfmsgllvgsflwgllysfgrfkpmalswccqlavagstlfpactfvlcyqlr 175
QY 170 ----- 169
DB 176 fvaafgmagiflsjllmewettsterravtmvvgcafsaggaalglaialrdwrtgl 235
QY 170 -----WLVESARWLITTNKLDDEGLKALRKVARTNGTKNAE----- 204
DB 236 aasvpfaaislswlipesarwllkqkpdqalqelrkvaringhkeaknlltletppppp 295
QY 205 ----- 204
DB 296 plpipsptapjstpcitftaltspappipseltpgpplqnisttfttvlignstll 355
QY 205 -----ETLNIEVNSTMOBELDAOTKTITVCD 231
DB 356 insthvtstlthlhthvarthimgvldvsgetecvyllkvlmsvkeevaseakepsvld 415
QY 232 LFRNPSMRKRKICLIVFLR 249
DB 416 lfcpvplwrtscamlvkvk 433

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 11:36:34 ; Search time 13.05 Seconds
(without alignments)
501.614 Million cell updates/sec

Title: US-09-674-235-1

Perfect score: 1406
Sequence: 1 MAFELLISQVGLGRFQMLH.....RKIKSRKHKNDCTKTKF 268

Scoring table: BLOSUM62
Gapop 10.0 , Capext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents-AA: *
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep: *
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep: *
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep: *
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep: *
5: /cgn2_6/prodata/2/1aa/PCBUS.COMB.pep: *
6: /cgn2_6/prodata/2/1aa/backfile1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	367.5	26.1	537	2	US-08-647-397-2 Sequence 2, Appli
2	169.5	12.1	536	3	US-08-501-572-1 Sequence 1, Appli
3	169.5	12.1	556	3	US-09-040-444-1 Sequence 1, Appli
4	152.5	10.8	555	3	US-08-501-572-3 Sequence 3, Appli
5	152.5	10.8	555	3	US-09-040-444-3 Sequence 3, Appli
6	146	10.4	553	3	US-08-501-572-2 Sequence 2, Appli
7	146	10.4	553	3	US-09-040-444-2 Sequence 2, Appli
8	104	7.4	520	4	US-08-964-127-2 Sequence 2, Appli
9	104	7.4	520	4	US-09-496-692-2 Sequence 2, Appli
10	100	7.1	524	2	US-08-928-692-12 Sequence 4, Appli
11	98.5	7.0	286	4	US-08-964-127-4 Sequence 4, Appli
12	98.5	7.0	286	4	US-09-496-692-4 Sequence 4, Appli
13	87	6.2	581	2	US-08-989-386-7 Sequence 7, Appli
14	83.5	5.9	494	2	US-09-031-392-5 Sequence 5, Appli
15	83.5	5.9	494	4	US-09-299-549-5 Sequence 5, Appli
16	83.5	5.9	494	4	US-09-610-417-5 Sequence 5, Appli
17	82	5.8	492	4	US-08-355-844-3 Sequence 3, Appli
18	82	5.8	492	5	PCT-US95-16126-3 Sequence 3, Appli
19	80.5	5.7	325	2	US-09-055-097-5 Sequence 5, Appli
20	80.5	5.7	521	1	US-08-075-193-2 Sequence 2, Appli
21	80.5	5.7	521	2	US-08-564-090A-2 Sequence 2, Appli
22	80.5	5.7	521	5	PCT-US94-06698-2 Sequence 2, Appli
23	80.5	5.7	1308	2	US-08-996-644-2 Sequence 2, Appli
24	80.5	5.7	1308	2	US-09-352-552-2 Sequence 2, Appli
25	79	5.6	285	3	US-09-187-050-12 Sequence 12, Appli
26	79	5.6	393	3	US-09-187-050-2 Sequence 2, Appli
27	79	5.6	393	3	US-09-187-050-14 Sequence 14, Appli

28	79	5.6	393	3	US-09-187-050-16 Sequence 16, Appli
29	79	5.6	393	3	US-09-187-050-18 Sequence 18, Appli
30	79	5.6	393	3	US-09-187-050-20 Sequence 20, Appli
31	79	5.6	393	3	US-09-187-050-22 Sequence 22, Appli
32	79	5.6	393	3	US-09-187-050-24 Sequence 24, Appli
33	79	5.6	393	3	US-09-187-050-26 Sequence 26, Appli
34	79	5.6	393	3	US-09-187-050-28 Sequence 28, Appli
35	79	5.6	393	3	US-09-187-050-29 Sequence 29, Appli
36	79	5.6	393	3	US-09-187-050-30 Sequence 30, Appli
37	79	5.6	393	3	US-09-187-050-31 Sequence 31, Appli
38	79	5.6	393	3	US-09-187-050-32 Sequence 32, Appli
39	79	5.6	393	3	US-09-187-050-33 Sequence 33, Appli
40	79	5.6	393	3	US-09-187-050-34 Sequence 34, Appli
41	79	5.6	393	3	US-09-187-050-34 Sequence 34, Appli
42	77	5.5	1189	4	US-09-176-320-8 Sequence 8, Appli
43	76.5	5.4	1642	2	US-08-447-411-45 Sequence 45, Appli
44	76.5	5.4	1642	2	US-08-662-227-2 Sequence 2, Appli
45	76.5	5.4	1642	4	US-09-017-947-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-647-397-2
Sequence 2, Application US/08647397
Patent No. 5972702

GENERAL INFORMATION:

APPLICANT: Beier, David R.

TITLE OF INVENTION: OSTEOCLAST TRANSPORTER

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Wolf, Greenfield & Sacks, P.C.

STREET: 600 Atlantic Avenue

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/647,397

FILING DATE:

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Gates, Edward R.

REGISTRATION NUMBER: 31,616

REFERENCE/DOCKET NUMBER: B0801/7048

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-720-2441

TELEFAX: 617-720-2441

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 537 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-647-397-2

Query Match 26.1%; Score 367.5; DB 2; Length 537;

Best local Similarity 24.6%; Pred. No. 1.2e-33;

Matches 88; Conservative 45; Mismatches 74; Indels 151; Gaps 5;

QY 1 MAFELLISQVGLGRFQMLH...IPILSMLLIPHLLNFPAALPGHRC-----W 50

Db 1 MTFSEILDRVSGMPQYHVTLLALPILGIANHNLQITATTPDHCKRPPNASLEPW 60

QY 51 VHLMDNNTGSGNTEGLSEDAALLRISIPIDSNLRPEKCRFPVHPQWOLHLNGTISTSE 110

Db 61 V-----LPLDPNGKPEKCLRFVHLP-----NMSLPNDTQ 89
Qy 111 ADHEPCVDGVNVOQSYFPSTIYTKMDLYCDYOSIKSVYVQFLLTGLMVGIGIGVSDR- 169
Db 90 GATEPCLDGMWYNT--RDITVEMDLYCGSNKLEMAQSVFMAGILVGPVPGELSDRF 147
Qy 170 -----169
Db 148 GRRPLTWGTYLLLAAGSSAAPSPLVYVMIFFELCGCSISGISLSTILLNVEWVPTSTR 207
Qy 170 -----WLVESARMLITN 182
Db 208 AASSTTTCYITGFIPLGLAVAPQWRMLQLVSAAFFISLMSWVPESTRMLVLSG 267
Qy 183 KLDGLKALRKVARTNGIKNAEETINIEVVRSTMOEELDAQTKTVCLFRNPSMRK 240
Db 268 KSRALKTLQRTATNGKKEBEKLTVEELKFNLDKDTISAKVYKGLSDLFRVSLIR 325

RESULT 2

US-08-501-572-1
Sequence 1, Application US/08501572
Patent No. 6063623
GENERAL INFORMATION:
APPLICANT: Koepsell, Hermann
APPLICANT: Grundeman, Dirk
APPLICANT: Gorboulev, Valentin
TITLE OF INVENTION: Transport protein which effects the
TITLE OF INVENTION: Transport of cationic xenobiotics and/or pharmaceuticals,
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/501,572
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Noohy, Kimberlin M
REGISTRATION NUMBER: 35,391
REFERENCE/DOCKET NUMBER: 02481.1453-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)408-4000
TELEFAX: (202)408-4400
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 556 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-501-572-1

Query Match 12.1%; Score 169.5; DB 3; Length 556;
Best Local Similarity 20.5%; Pred. No. 7.8e-11;
Matches 77; Conservative 47; Mismatches 95; Indels 157; Gaps 15;

Qy 4 EELISQVGGGLRFO---MLHVLFLPSLMLLPHILLENFAAIPGHC-----49
Db 5 DDVLEQVGEFQKQAFLLILCLISASLAPIYGVIGLFT---PGHYCQNPVAELSOR 61

Qy 50 --WYHMLDN-----TSGNETGILSBDALLRISIPDSNLREPKCRFRVHPQMLHLN 102
Db 62 CGNSQAEELNYVYGLDPSSEAFSLSO--CMRYEV--DMNOSTLDC---VDP-----LS 108
Qy 103 GIHTSTSEADTEPCVDGVNVOQSYFPSTIYTKMDLYCDYOSIKSVYVQFLLTGLMVGIT 162
Db 109 SLVANSRLPLPGCEHGMWYDTP--GSSITYERNLVCGDMKVDLFQSCVNLFFELGSLV 166
Qy 163 GGHVSDR-----169
Db 167 VGTIADPFGKLLCLLVTLTVSVSVGLTAVAPDYTNLFRLLQGMVSKGSWVGYTLIT 226
Qy 170 -----WLV-ES 174
Db 227 EFVSGYRRTTALTYQMAFTVGLVGLAGVAYALPDWRMLQLAVSLPTFLLLYWVFPES 286
Qy 175 ARRLITNKLDEGLKALRKVARTNG--IKNAEETINIEVVRSTMOEELDAQTKT-TVCDL 232
Db 287 FHWLSQKRTTAVRIMEDIAQKNGKVPAD-----LKMCLDE--DASEKRSPEFADL 338
Qy 233 FRNPSMRKRICILVFL 248
Db 339 FRPMLRKHTVLMYL 354

RESULT 3

US-09-040-444-1
Sequence 1, Application US/09040444
Patent No. 6063766
GENERAL INFORMATION:
APPLICANT: Koepsell, Hermann
APPLICANT: Grundeman, Dirk
APPLICANT: Gorboulev, Valentin
TITLE OF INVENTION: Transport protein which effects the
TITLE OF INVENTION: Transport of cationic xenobiotics and/or pharmaceuticals,
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner, L.L.P.
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/040,444
FILING DATE: March 18, 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: O'Connor, Steven P
REGISTRATION NUMBER: 41,225
REFERENCE/DOCKET NUMBER: 2481.1453-01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)408-4000
TELEFAX: (202)408-4400
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 556 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-040-444-1

Query Match 12.1%; Score 169.5; DB 3; Length 556;
Best Local Similarity 20.5%; Pred. No. 7.8e-11;
Matches 77; Conservative 47; Mismatches 95; Indels 157; Gaps 15;

RESULT 4
 US-08-501-572-3, Application US/08501572
 : Sequence 3, Patent No. 6063623
 : GENERAL INFORMATION:
 : APPLICANT: Koepsell, Herman
 : APPLICANT: Grubbe, Valentin
 : APPLICANT: Grubbe, Valentin
 : TITLE OF INVENTION: Transport of Cationic Xenobiotics and/or Pharmaceuticals
 : TITLE OF INVENTION: Transport of Cationic Xenobiotics and/or Pharmaceuticals
 : TITLE OF INVENTION: DNA Sequences Encoding It and Their Use.
 : NUMBER OF SEQUENCES: 6
 : CORRESPONDENCE ADDRESS: Finnegun, Henderson, Farabow, Garrett & Dunner
 : ADDRESS: 1300 I Street, N.W., Suite 700
 : STREET: Washington
 : CITY: Washington
 : STATE: D.C.
 : COUNTRY: USA
 : ZIP: 20005-3315
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent in Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/501,572
 : FILING DATE:
 : CLASSIFICATION: 424
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Toohey, Kimberlin M
 : REGISTRATION NUMBER: 35,391
 : REFERENCE/DOCKET NUMBER: 02481.1453-00000
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (202)408-4400
 : TELEFAX: (202)408-4400
 : INFORMATION FOR SEQ ID NO: 3:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 555 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: peptide
 : US-08-501-572-3

[illegible]

RESULT 5
 US-09-040-444-3
 Sequence 3, Application US/09040444
 Patent No. 6063766
 GENERAL INFORMATION:
 APPLICANT: Koepsell, Hermann
 APPLICANT: Grudeman, Dirk
 APPLICANT: Gorbulev, Valentin
 TITLE OF INVENTION: Transport Of Cationic Xenobiotics and/or Pharmaceuticals
 TITLE OF INVENTION: Transport Of Cationic Xenobiotics and Their Use.
 TITLE OF INVENTION: DNA Sequences Encoding It And Their Use.
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner, L.L.P.
 STREET: 1300 I Street, N.W., Suite 700
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005-3315
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER SYSTEM: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION NUMBER: US/09/040,444
 APPLICATION FILING DATE: March 18, 1998
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: O'Connor, Steven P
 REGISTRATION NUMBER: 41,225
 REFERENCE/DOCKET NUMBER: 2481.1453-01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202)408-4400
 TELEFAX: (202)408-4400
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 555 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide

Tue Jul 2 10:13:39 2002

US-09-040-444-3

Query Match
Best Local Similarity 10.8%; Score 152.5; DB 3; Length 555;
Matches 71; Conservative 40; Mismatches 109; Indels 139; Gaps 12;

QY 11 GGLGRFQMLHLVFLIPSLMLLPIHLLE-NPAAIPGHRG---VWMLDNNNGSGNGEIGI 66
DB 12 GGEFFRQKQFLLALSLAFATVYGIYLFETPDHCRSPGVALSLRG-----64
QY 67 LSEALLRISIP---LDSNLRPEKRRFVHPQQLHLNGTHTSSEADTF-----PC 116
DB 65 WSPAELNLYTVPGPGAGESPQCRRY-EVWMSQSTFD-CVPLASLDRSLPLGPC 122
QY 117 VDSMVYDOSYFSTYTKMDLVCDYQSLKSVQFLLTGMLVGITGGHVSF-----169
DB 123 RDGMVYEPF--GSSIVTEFNLVCANSMWMLDFQSSVANGFPGSGISGYIADRFGRKLC 180
QY 170 -----WLV-----172
DB 181 LFTVLAAGVLAISPPTWMLIFRLQGLVSKAGMLIGLITTEFYGRRTVCI 173
QY 173 -----ESARLITTKLDEGL 188
DB 241 YQAVYVGLVAVAYALPHRMLOFVALPNEFFLYWCIPSPFRLISQNKNEAM 300
QY 189 KALKRVARTNGIKNAEETLNIEVYSTMOEELDAQOTKTYCDLFRNSMRRICILVF 247
DB 301 RIHKIAKKNG-----KSLPASLQRLREE-TGKKLNPFLDLVKTPOIKRHTMILMY 353

RESULT 6
US-08-501-572-2
Sequence 2, Application US/08501572
Patent No. 6063623
GENERAL INFORMATION:
APPLICANT: Koepsell, Hermann
APPLICANT: Gorboulev, Valentin
TITLE OF INVENTION: Transport protein which effects the
TITLE OF INVENTION: Transport of Cationic Xenobiotics and/or Pharmaceuticals,
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/501.572
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Toohy, Kimberlin M
REGISTRATION NUMBER: 35,391
TELECOMMUNICATION INFORMATION: 02481.1453-00000
TELEPHONE: (202)408-4400
TELEFAX: (202)408-4400
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 553 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-09-674-235-1.ra1

MOLECULE TYPE: peptide
US-08-501-572-2

Query Match
Best Local Similarity 10.4%; Score 146; DB 3; Length 553;
Matches 75; Conservative 34; Mismatches 117; Indels 142; Gaps 11;

QY 4 EELLQVGLGFRQMLHLVFLIPSLMLLPIHLLENPAAIPGHRG---VWMLDNNNGS 60
DB 5 DDLQVGSQWQKQAFLLICLSAFAFCVGIYLFETPDHCRSPGVALSLRG-----63
QY 61 GNEGTISDALRISIPDSNLRPE-----KRRFVHPQQLHLN-----GTHT 108
DB 64 -----WSPAELNLYTVP---GLGPGAEFLGQCRRY-EVWMSQSTFD-CVPLASL 113
QY 109 SEADTEPCVDGWYVDQSYFSTYTKMDLVCDYQSLKSVQFLLTGMLVGITGGHVS 168
DB 114 SHLPLGRCQDQWYDIP--GSSIVTEFNLVCADSWKIDLPOSLNAGFFFGSLGVYF 171
QY 169 R-----169
DB 172 RGRKLCILGTVLVAVSGLVMAFSPVMSMLFRLQGLVSKGNMAGYTLITTEYSG 231
QY 170 -----WLV-ESARLITTKLDEGL 180
DB 232 SRRVIAIYQMAFTVGLVALGLAYALPHRMLOLAVSLPFLFLYWCIPSPFRLIS 291
QY 181 TNKLDGKALKRVARFNGIKNAEETLNIEVYSTMOEELDAQOTKTYCDLFRNSMRR 240
DB 292 QKRTTEAIKIMDIAQKNG-KLPPADLKMISLEEDYTEKL-----SPSEADLFRTPRL 345
QY 241 RICILVEL 248
DB 346 RTFLIMYL 353

RESULT 7

US-09-040-444-2
Sequence 2, Application US/09040444
Patent No. 6063766
GENERAL INFORMATION:
APPLICANT: Koepsell, Hermann
APPLICANT: Gorboulev, Valentin
TITLE OF INVENTION: Transport protein which effects the
TITLE OF INVENTION: Transport of Cationic Xenobiotics and/or Pharmaceuticals,
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finegan, Henderson, Farabow, Garrett & Dunner, L.L.P.
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/040.444
FILING DATE: March 18, 1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: O'Connor, Steven P
REGISTRATION NUMBER: 41,225
TELECOMMUNICATION INFORMATION: 2481.1453-01
TELEPHONE: (202)408-4000
TELEFAX: (202)408-4400
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 286 amino acids
TYPE: amino acid
TOPOLOGY: linear

7

MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-964-127-4

Query Match 7.0%; Score 98.5; DB 4; Length 286;
Best Local Similarity 17.8%; Pred. No. 0.0043;

Matches 54; Conservative 29; Mismatches 73; Indels 147; Gaps 11;

39 NFAAIPGHCWVMDLNTGSGNETGILSEDAIRISIPIDSRLPCKCRFPVHPOMQL 98
16 HYGAFPPNAGSG -EQPPNAGSVASALAAASRAVATSDPS-----CSGFAP----- 65
99 LHLNGTHTSHSEADTEPCVDGWYD -QSYFSTITVKMDVCD--YQSLKSVQFL--- 152
66 -----DENHCLKMDYNGLPVLTNNAGMDVCDLGMOVYILQILFILGFA 112
153 -----LTGMLVSGI----- 161
113 SGYFLGYPADRFGRGIVLLTGLVPGCVGGAAGSSTGVMALRFLGLAGVLDGV 172
162 -----IGGH-----VSDRW----- 170
173 YLMRELCDPTQRLRVVALAGELVGVGHFLFLGLALVSKDMRFLORMITAPCILEFLYGV 232
171 ---LYESARWLITNKLDGELKALKRYA---RTNGIKNAEETINIEVVRSTMOEELDAQ 224
233 PGLFLESARWLIVKRIEASQSVLRILAEARNRHGOMLGEA-----QEAOLDLE 282
225 TKT 227
283 SST 285

RESULT 12

US-09-496-692-4

Sequence 4, Application US/09496692

Patent No. 6313271

GENERAL INFORMATION:

APPLICANT: Grandearl, Andrew David John

TITLE OF INVENTION: NOVEL GENES ENCODING TRANSPORTER-LIKE

TITLE OF INVENTION: MOLECULES

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows 95

SOFTWARE: FastSeq for Windows Version 2.0b

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/496,692

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/964,127

FILING DATE: 06-NOV-1997

ATTORNEY/AGENT INFORMATION:

NAME: Crews, Ph.D., L. Lee

REGISTRATION NUMBER: P-43,567

REFERENCE/DOCKET NUMBER: 07334/038001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 286 amino acids

type: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-09-496-692-4

Query Match 7.0%; Score 98.5; DB 4; Length 286;
Best Local Similarity 17.8%; Pred. No. 0.0043;

Matches 54; Conservative 29; Mismatches 73; Indels 147; Gaps 11;

39 NFAAIPGHCWVMDLNTGSGNETGILSEDAIRISIPIDSRLPCKCRFPVHPOMQL 98
16 HYGAFPPNAGSG -EQPPNAGSVASALAAASRAVATSDPS-----CSGFAP----- 65
99 LHLNGTHTSHSEADTEPCVDGWYD -QSYFSTITVKMDVCD--YQSLKSVQFL--- 152
66 -----DENHCLKMDYNGLPVLTNNAGMDVCDLGMOVYILQILFILGFA 112
153 -----LTGMLVSGI----- 161
113 SGYFLGYPADRFGRGIVLLTGLVPGCVGGAAGSSTGVMALRFLGLAGVLDGV 172
162 -----IGGH-----VSDRW----- 170
173 YLMRELCDPTQRLRVVALAGELVGVGHFLFLGLALVSKDMRFLORMITAPCILEFLYGV 232
171 ---LYESARWLITNKLDGELKALKRYA---RTNGIKNAEETINIEVVRSTMOEELDAQ 224
233 PGLFLESARWLIVKRIEASQSVLRILAEARNRHGOMLGEA-----QEAOLDLE 282
225 TKT 227
283 SST 285

RESULT 13

US-08-989-386-7

Sequence 7, Application US/08989386

Patent No. 5989860

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Hillman, Jennifer L.

APPLICANT: Guejter, Karl J.

APPLICANT: Corley, Neil C.

APPLICANT: Shah, Puryi

TITLE OF INVENTION: HUMAN ISOMERASE HOMOLOGS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/989,386

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0443 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

[illegible]

Tue Jul 2 10:13:39 2002

us-09-674-235-1.rail

Page 9

Search completed: July 1, 2002, 11:39:26
job time: 172 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 11:39:30 ; Search time 11.68 seconds
(without alignments)
888.428 Million cell updates/sec

Title: US-09-674-235-1

Perfect score: 1406
Sequence: 1 MAFEEILSGVGLGRFQMLH.....KKIKSRKHKNDCTTKVTKF 268

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	216.5	15.4	557	1 OCN2_MOUSE	092068 mus musculus
2	195.5	13.9	557	1 OCN2_RAT	070594 rattus norv
3	193.5	13.8	557	1 OCN2_HUMAN	076082 mus sapien
4	100	7.1	524	1 GTR2_HUMAN	P11168 homo sapien
5	95	6.8	523	1 GTR2_MOUSE	P14246 mus musculu
6	93.5	6.7	526	1 C11A_MOUSE	094882 mus musculu
7	93	6.6	522	1 GTR2_RAT	P12336 rattus norv
8	87	6.2	581	1 FKX_MOUSE	061576 mus musculu
9	86.5	6.2	496	1 GTR3_CHICK	P28568 gallus gall
10	84.5	6.0	636	1 GYRB_THEMA	P77993 thermocoga
11	83.5	5.9	494	1 GTR3_BOVIN	P58332 bos taurus
12	83.5	5.9	494	1 GTR3_SHEEP	P47843 ovis aries
13	83.5	5.9	520	1 C11A_CAPRI	P79183 capra hircu
14	82.5	5.9	490	1 GTR1_CHICK	P46896 gallus gall
15	82.5	5.9	555	1 GTR3_SCHPO	092339 schizosacch
16	82	5.8	492	1 GTR1_HUMAN	P11166 homo sapien
17	81.5	5.8	400	1 GTR3_RABIT	Q94392 oryctolagus
18	81.5	5.8	608	1 YD56_YEAST	Q93822 saccharomyc
19	81	5.8	520	1 C11A_PIG	P10612 sus scrofa
20	81	5.8	561	1 YGGA_YEAST	P53166 saccharomyc
21	80.5	5.7	260	1 CDH_MYCTU	050676 mycobacteri
22	80.5	5.7	521	1 C11A_HUMAN	P49071 drosophila
23	80	5.7	359	1 NKK2_DROME	P17809 mus musculu
24	80	5.7	492	1 GTR1_MOUSE	P13355 oryctolagus
25	80	5.7	492	1 GTR1_RABIT	P11167 rattus norv
26	80	5.7	492	1 GTR1_RAT	Q94814 mesocricetu
27	80	5.7	520	1 C11A_MESAU	009314 caenorhabdi
28	80	5.7	757	1 YOF2_CAEEL	046515 equus caball
29	79.5	5.7	520	1 C11A_HORSE	P13141 homo sapien
30	79	5.6	198	1 SDC4_HUMAN	Q10503 mycobacteri
31	79	5.6	364	1 YW1_MYCTU	P00189 bos taurus
32	79	5.6	520	1 C11A_BOVIN	Q14651 homo sapien
33	79	5.6	629	1 PLST_HUMAN	

34	78.5	5.6	495	1 GTR3_CANFA	P47842 canis famli
35	78	5.5	1148	1 A8A2_MOUSE	P98200 mus musculu
36	77.5	5.5	1124	1 PHX_MOUSE	P33523 mougeotia s
37	77	5.5	270	1 VSP1_ARATH	O49195 arabidopsis
38	77	5.5	526	1 C11A_RAT	P14137 rattus norv
39	77	5.5	615	1 SPRA_HAERIN	P45243 haemophilus
40	77	5.5	1693	1 R1P2_MOUSE	P97433 mus musculu
41	76.5	5.4	201	1 YPT1_PHTIN	001890 phytophthor
42	76.5	5.4	496	1 GTR3_HUMAN	P11169 homo sapien
43	76.5	5.4	1420	1 APX_XENLA	Q01613 xenopus lae
44	76	5.4	315	1 VN34_ROT54	Q00721 stigmatia 11 r
45	76	5.4	376	1 Y433_MYCTU	P96279 mycobacteri

ALIGNMENTS

```

RESULT 1
OCN2_MOUSE STANDARD; PRT: 557 AA.
ID OCN2_MOUSE
AC 092068;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Organic cation/carnitine transporter 2 (Solute carrier family 22, member 5) (High-affinity sodium-dependent carnitine cotransporter).
GN SLC22A5 OR OCTN2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OC NCBI_Taxid:10090;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Kidney;
RC MEDLINE=99113835; Pubmed=9916797;
RA Nezu J., Tamai I., Oku A., Ohashi R., Yabuuchi H., Hashimoto N.,
RA Nishikido H., Sai Y., Koizumi A., Shoji Y., Takada G., Matsushita T.,
RA Yashino M., Kato H., Ohura T., Tsujimoto G., Hayakawa J., Shimane M.,
RA Tsuji A.;
RT "Primary systemic carnitine deficiency is caused by mutations in a
RT gene encoding sodium ion-dependent carnitine transporter.";
RT Nat. Genet. 21:91-94(1999).
[2]
RN SEQUENCE FROM N.A., AND VARIANT JVS ARG-352.
RP STRAIN=C3H;
RC MEDLINE=99057546; Pubmed=9837751;
RA Lu K., Nishimori H., Nakamura Y., Shima K., Kuwajima M.;
RT "A missense mutation of mouse OCTN2, a sodium-dependent carnitine
RT cotransporter, in the juvenile visceral steatosis mouse.";
RT Biochem. Biophys. Res. Commun. 252:590-594(1998).
[3]
RN SEQUENCE FROM N.A., AND CHARACTERIZATION.
RP MEDLINE=99384224; Pubmed=10454528;
RA Wu X., Huang W., Prasad P.D., Seth P., Rajan D.P., Leibach F.H.,
RA Chen J., Conway S.J., Ganapathy V.;
RT "Functional characteristics and tissue distribution pattern of organic
RT cation transporter 2 (OCTN2), an organic cation/carnitine
RT transporter.";
RT J. Pharmacol. Exp. Ther. 290:1482-1492(1999).
[4]
RN FUNCTION: SODIUM-ION DEPENDENT, HIGH AFFINITY CARNITINE
RN TRANSPORTER. ALSO TRANSPORTS ORGANIC CATIONS WITHOUT THE
RN INVOLVEMENT OF SODIUM. INVOLVED IN THE ACTIVE CELLULAR UPTAKE OF
RN CARNITINE.
RN SUBCELLULAR LOCATION: Integral membrane protein.
RN DISEASE: DEFECTS IN SLC22A5 ARE THE CAUSE OF JUVENILE VISCERAL
RN STEATOSIS (JVS).
RN SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. ORGANIC
RN CATION SUBFAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way

```

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: AB015800; BAA36590.1; -
 DR EMBL: AF111425; AAC9787.1; -
 DR EMBL: AF110417; AAD54060.1; -
 DR MGI: 1329012; S1c22a5.
 DR InterPro: IPR003662; sub-transporter.
 DR Pfam: PF00083; sugar_tr.1.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
 KW Transport; Transmembrane; Glycoprotein; Disease mutation.
 FT TRANSMEM 21 41
 FT TRANSMEM 143 163
 FT TRANSMEM 173 193
 FT TRANSMEM 198 218
 FT TRANSMEM 233 253
 FT TRANSMEM 258 278
 FT TRANSMEM 342 362
 FT TRANSMEM 374 394
 FT TRANSMEM 407 427
 FT TRANSMEM 431 451
 FT TRANSMEM 489 509
 FT CARBOHYD 57 64
 FT CARBOHYD 91 91
 FT CARBOHYD 322 322
 FT CARBOHYD 352 352
 FT VARIANT 352 352 L -> R (IN JVS).
 SQ SEQUENCE 557 AA: 62779 MW: 6093f0EB9612B204 CRC64;

Query Match 15.4%; Score 216.5; DB 1; Length 557;
 Best Local Similarity 23.4%; Pred. No. 1.2e-11;
 Matches 91; Conservative 34; Mismatches 89; Indels 175; Gaps 14;

QY 3 FEELLSQVGLGRFOMHLVFLPMLLPH-----ILENFAAIPGRGV-HMLD 55
 DB 4 YDEVTAFLGEMGPFQ--RLTFLLSAS-IIPNGFNGMSIV--FLAGTPHRCVPTIVN 57
 QY 56 NNTSGNENGLSEDLALRLISIP---DSNLREKCRFPHPQMLLNGTSTSEAD 112
 DB 58 LSSAVNNH-----SIPLETKDGRQVQKCRRY-----RLATIANSELG 96
 QY 113 TEP-----CVNGWYDOSYFPTSTYTKMDVCDYQSLKSVQFLLTGMVG 159
 DB 97 LEPGRDVLEQLQESCLDGMEDKDVFLSTVTEMDLVCKDKMKAPLITSLFPGVLMG 156
 QY 160 GILGHVSDR----- 169
 DB 157 SPISGOLSDRFGRKNVFLITMGQTFSEFLQVSVNEFEFTVLVGVGMQISNVYAFV 216
 QY 170 ----- 171
 DB 217 LGTEILSKSIRLIFATVLCIPYAFGFMVLPFAFYFIRDMRMLLALVPGVLCALMWF 276
 QY 172 V-ESARMLITNKLEDEGKALKRKVARINGIKRAEETLNEVVRSTM---QEELDAAGTKT 227
 DB 277 IPESPRMILSGRIEAEVITIRKAKINGI-----VAPSTIIFPSELQDLNSTKP 326
 QY 228 T---VCDLFNPSMKRKLICILVLRKKIS 253
 DB 327 QLHHITDILFRIRIVITIMSIILMTIS 355

RESULT 2
 OCN2_RAT STANDARD; PRT; 557 AA.
 AC 070594; 09QWL0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Organic cation/carnitine transporter 2 (Solute carrier family 22, member 5) (High-affinity sodium-dependent carnitine cotransporter)

DE (US2r) (Ct1).
 GN SUC22A5 OR OCTN2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA MEDLINE=98200080; PubMed=9541011;
 RA Schoenly E., Spitznberger F., Engelhardt M., Martel F., Oeding N.,
 RA Gruendemann D.;
 RT "Molecular cloning and characterization of two novel transport
 RL proteins from rat kidney";
 RN FEBS Lett. 425:79-86(1998).
 RP [2]
 RC SEQUENCE FROM N.A.
 RA STRAIN=SPRAGUE-DAWLEY; TISSUE=Intestine;
 RA MEDLINE=99011422; PubMed=9792817;
 RA Sekine T., Kusuhara H., Utsunomiya-Tate N., Tsuda M., Sugiyama Y.,
 RA Kanai Y., Endou H.;
 RT "Molecular cloning and characterization of high-affinity carnitine
 RL transporter from rat intestine";
 RN Biochem. Biophys. Res. Commun. 251:586-591(1998).
 RP [3]
 RC SEQUENCE FROM N.A. AND CHARACTERIZATION.
 RA MEDLINE=99384224; PubMed=10454528;
 RA Wu X., Huang W., Prasad P.D., Seth P., Rajan D.P., Leibach F.H.,
 RA Chen J., Conway S.J., Ganapathy V.;
 RT "Functional characteristics and tissue distribution pattern of organic
 RL cation transporter 2 (OCTN2), an organic cation/carnitine
 RT transporter";
 RC J. Pharmacol. Exp. Ther. 290:1482-1492(1999).
 CC -1- FUNCTION: SODIUM-ION DEPENDENT, HIGH AFFINITY CARNITINE
 CC TRANSPORTER. ALSO TRANSPORTS ORGANIC CATIONS WITHOUT THE
 CC INVOLVEMENT OF SODIUM. INVOLVED IN THE ACTIVE CELLULAR UPTAKE OF
 CC CARNITINE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC AND IN THE GLOMERULI IN THE KIDNEY, IN THE MYOCARDIUM, VALVES, AND
 CC ARTERIOLES IN THE HEART, IN THE LABYRINTHINE LAYER OF THE
 CC PLACENTA, AND IN THE CORTEX, HIPPOCAMPUS, AND CEREBELLUM IN THE
 CC BRAIN.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. ORGANIC
 CC CATION SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).


```

QY 170 ----- 169
Db 169 RKNVLFYTMGOTGFSFLQIFSKNFMFVFLVGMGQISNVAFAVLGTETLLKRSYRI 228
QY 170 ----- -WLV-ESARPLIITN 182
Db 229 IFSTGVCIEFAFGYMWVLFPAFYIRDWRMVLVLTMPGVLCAVLMWVIFPESPRLISOG 288
QY 183 KIDEELKMLRKVARFKARKNGKMAEETLINEVNST-----MOEELDAOATKTYVDLFENP 236
Db 289 RFEELVETIRRAAKANGI-----VVESTIFDPSELQDLSSKKQSHNIIDLRTW 338
QY 237 SMKRRCILVFLRKKS 253
Db 339 NIRMVYTIMSLMMTIS 355

RESULT 4
GTR2_HUMAN
ID GTR2_HUMAN STANDARD: PRT; 524 AA.
AC P1168;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Solute carrier family 2, facilitated glucose transporter, member 2
DE (Glucose transporter type 2, liver).
DE SLC2A2 OR GLUT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eulambda; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANT LEU-68.
RX TISSUE=Liver, and Kidney;
RX MEDLINE=88289735; PubMed=3399500;
RA Fukunoto H., Seino S., Imura H., Steno Y., Eddy R.L., Fukushima Y.,
RA Byers M.G., Shown T.B., Bell G.I.,
RT "Sequence, tissue distribution, and chromosomal localization of mRNA
RT encoding a human glucose transporter-like protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:5434-5438(1988).
CC -1- FUNCTION: FACILITATING GLUCOSE TRANSPORTER. THIS ISOPFORM LIKELY
CC MEDIATES THE BIDIRECTIONAL TRANSFER OF GLUCOSE ACROSS THE PLASMA
CC MEMBRANE OF HEPATOCYTES AND IS RESPONSIBLE FOR UPTAKE OF GLUCOSE
CC BY THE BETA CELLS; MAY COMPRISE PART OF THE GLUCOSE-SENSING
CC MECHANISM OF THE BETA CELL. MAY ALSO PARTICIPATE WITH THE
CC NA(+)/GLUCOSE COTRANSPORTER IN THE TRANSCELLULAR TRANSPORT OF
CC GLUCOSE IN THE SMALL INTESTINE AND KIDNEY.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: LIVER, INSULIN-PRODUCING BETA CELL, SMALL
CC INTESTINE, AND KIDNEY.
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
CC TRANSPORTER SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL, J03810; AAA59514.1; -.
DR PIR, A1318; A31318.
DR MIM, 138160; -.
DR InterPro: IPR0033663; Sugar_transporter.
DR InterPro: IPR0033662; sub_transporter.
DR Pfam, PF00083; sugar_tr; 1.
DR PRINTS, PR00171; SUGRTNSPORT.
DR PROSITE, PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE, PS00217; SUGAR_TRANSPORT_2; 1.
DR Transmembrane, Sugar transporter; Transport; Glycoprotein;
DR Multiene family; polyomorphism

```

```

FT DOMAIN 1 10 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 11 31 1 (POTENTIAL).
FT DOMAIN 32 98 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 99 119 2 (POTENTIAL).
FT DOMAIN 120 127 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 128 148 3 (POTENTIAL).
FT DOMAIN 149 158 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 159 175 4 (POTENTIAL).
FT DOMAIN 180 187 5 (POTENTIAL).
FT TRANSMEM 188 208 6 (POTENTIAL).
FT DOMAIN 209 217 6 (POTENTIAL).
FT TRANSMEM 218 238 7 (POTENTIAL).
FT DOMAIN 239 303 7 (POTENTIAL).
FT TRANSMEM 304 324 8 (POTENTIAL).
FT DOMAIN 325 338 8 (POTENTIAL).
FT TRANSMEM 339 359 9 (POTENTIAL).
FT DOMAIN 360 368 9 (POTENTIAL).
FT TRANSMEM 369 389 10 (POTENTIAL).
FT DOMAIN 390 400 10 (POTENTIAL).
FT TRANSMEM 401 421 11 (POTENTIAL).
FT DOMAIN 422 433 11 (POTENTIAL).
FT TRANSMEM 434 454 12 (POTENTIAL).
FT DOMAIN 455 461 12 (POTENTIAL).
FT TRANSMEM 462 482 12 (POTENTIAL).
FT DOMAIN 483 524 12 (POTENTIAL).
FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 68 68 P -> L.
FT SEQUENCE 524 AA: 57489 MW: DA600577207EC083 CRC64: /FTid=VAR.007169.
SO Query Match 7.1%; Score 100; DB 1; Length 524;
Best Local Similarity 27.6%; Pred. No. 0.23;
Matches 35; Conservative 22; Mismatches 34; Indels 36; Gaps 6;

```

```

RA Asano T., Shibasaki Y., Iin J.L., Akanuma Y., Takaku F., Oka Y.;
RT "The nucleotide sequence of cDNA for a mouse liver-type glucose
RT transporter protein."
RL Nucleic Acids Res. 17:6386-6386(1989).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Liver;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batova S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochaya H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schiraldi L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barish G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamlaya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszewski B., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [4]
RP SEQUENCE OF 384-496 FROM N.A.
RC MEDLINE=92111400; PubMed=1165007;
RA Hogan A., Heyner S., Charron M.J., Copeland N.G., Gilbert D.J.,
RA Jenkins N.A., Thorens B., Schultz G.A.;
RT "Glucose transporter gene expression in early mouse embryos."
RL Development 113:363-372(1991).
RN [5]
RP TISSUE SPECIFICITY.
RC STRAIN=C57BL/6;
RX MEDLINE=93170163; PubMed=1289053;
RA Smith D.E., Gridley T.;
RT "Differential screening of a PCR-generated mouse embryo cDNA library:
RT glucose transporters are differentially expressed in early
RT postimplantation mouse embryos."
RL Development 116:555-561(1992).
CC -1- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. THIS ISOFORM LIKELY
CC MEDIATES THE BIDIRECTIONAL TRANSFER OF GLUCOSE ACROSS THE PLASMA
CC MEMBRANE OF HEPATOCYTES AND IS RESPONSIBLE FOR UPTAKE OF GLUCOSE
CC BY THE BETA CELLS. MAY COMPRISE PART OF THE GLUCOSE-SENSING
CC MECHANISM OF THE BETA CELL. MAY ALSO PARTICIPATE WITH THE
CC NA(+)/GLUCOSE COTRANSPORTER IN THE TRANSCELLULAR TRANSPORT OF
CC GLUCOSE IN THE SMALL INTESTINE AND KIDNEY.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: IN EMBRYO, EXPRESSED IN ENDODERM LAYER OF
CC YOLK SAC AND LIVER PRIMORDIUM.
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
CC TRANSPORTER SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X16986; CA34855.1; -
DR EMBL: X15684; CA33719.1; -
DR EMBL: S77926; AAB20847.1; -
DR EMBL: AK005068; BAB23792.1; -
DR PIR: S06920; S06920.
DR PIR: S05319; S05319.
DR MGD: MGI:1095438; Slc2a2.
DR InterPro: IPR003663; Sugar_transporter.

```

DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 DR PRINTS: PR00171; SUGRTNSPORT.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transmembrane: Sugar transport; Transport; Glycoprotein;
 MW Multigene family.
 FT DOMAIN 1 10
 FT TRANSSEM 11 31
 FT DOMAIN 32 97
 FT TRANSSEM 98 118
 FT DOMAIN 119 126
 FT TRANSSEM 127 147
 FT DOMAIN 148 157
 FT TRANSSEM 158 178
 FT DOMAIN 179 186
 FT TRANSSEM 187 207
 FT DOMAIN 208 216
 FT TRANSSEM 217 237
 FT DOMAIN 238 302
 FT TRANSSEM 303 323
 FT DOMAIN 324 337
 FT TRANSSEM 338 358
 FT DOMAIN 359 367
 FT TRANSSEM 368 388
 FT DOMAIN 389 401
 FT TRANSSEM 402 422
 FT DOMAIN 423 432
 FT TRANSSEM 433 453
 FT DOMAIN 454 460
 FT TRANSSEM 461 481
 FT DOMAIN 482 523
 FT CARBOHYD 62 62
 FT CONFLICT 106 106
 FT CONFLICT 181 181
 FT CONFLICT 327 327
 FT CONFLICT 351 351
 FT CONFLICT 432 432
 SO SEQUENCE 523 AA; 57106 MW; F12B5376CBDB5F63 CRC64;

Query Match
 Best Local Similarity 6.8%; Score 95; DB 1; Length 523;
 Matches 34; Conservative 25.6%; Pred. No. 0.65; Mismatches 36; Indels 36; Gaps 6;

QY 143 SLKSVQVFLTLTGMVGGITG-----CHVSDRWLV-----ESARW 177
 DB 186 ALGTLHQLALVTGILISQAGLSFLLGN-ODWHIILGLSAVPAALQCLLLECEPSRY 244
 QY 178 LIITNKLDGLKALKRVARTNGIKNAEETLINEVVRSTMQEELDAQT-KTYVCDLFRN 235
 DB 245 LVI--KLEEVRAKKSLKRLGTEDVTKDIN-----EMKKEKEASTEOKVSVIOLFTD 296
 QY 236 PSNRKRICLIVFL 248
 DB 297 ANYRQPIILVALML 309

RESULT 6
 C11A_MOUSE STANDARD: PRT; 526 AA.
 ID C11A_MOUSE
 AC Q9Q282;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Cytochrome P450 11A1, mitochondrial precursor (BC 1.14.15.6) (CYP11A1) (P450(scc)) (Cholesterol side-chain cleavage enzyme) (Cholesterol desmolase).
 GN CYP11A OR CYP11A1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RA Tanaka M., Hennebold J.H., Adashi E.Y.;
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CATALYZES THE SIDE-CHAIN CLEAVAGE REACTION OF
 CC CHOLESTEROL TO PREGNENOLONE (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Cholesterol + reduced adrenal ferredoxin +
 CC O(2) = pregnenolone + 4-methylpentanal + oxidized ferredoxin +
 CC ferredoxin + H(2)O.
 CC -1- PATHWAY: INITIAL RATE-LIMITING REACTION IN THE SYNTHESIS OF
 CC VARIOUS STEROID HORMONES.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF195119; AAF03897.1; -
 CC HSSP: P00189; 1SCC.
 DR MGD: MGI:88582; Cyp11a.
 DR InterPro: IPR001128; Cyt_P450.
 DR Pfam: PF00067; P450; 1.
 DR PRINTS: PR00385; P450.
 DR PROSITE: PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Membrane; Heme; Mitochondrion;
 KW Steroidogenesis; Transit peptide.
 FT TRANSIT 1 36
 FT CHAIN 37 526
 FT BINDING 458 458
 FT HEME (BY SIMILARITY).
 FT CYTOCHROME P450 11A1.
 SO SEQUENCE 526 AA; 60314 MW; E5029738EEAECB71 CRC64;

Query Match
 Best Local Similarity 6.7%; Score 93.5; DB 1; Length 526;
 Matches 55; Conservative 22.1%; Pred. No. 0.88; Mismatches 67; Indels 89; Gaps 13;

QY 51 VHMIDNNTGSGNETGILSEDLRLISIP-----LDSNLRPECRPVHQWQ 97
 DB 179 LHRRIKQNSGNSGVISD-LRFSESSISYIFGRMKLEIYDF-AQRFINAVYQ 236
 QY 98 LHLNFTI-----HST-----SEADTEPCVDGVYDSYFPSTIVT 133
 DB 237 MHTSVPMNLPPDFEFLRLRTKTKWKAAMVDYIFNKADL-----TYQNFY----- 282
 QY 134 KWDL--VCDYOSLSVVOFL-----LGMVLVGGIIGHVSDRW-LVESARW 177
 DB 283 -WDLRQKRDFSQYPGVLYSLGSKNKLPEKNIQANITEMLAGVDTSMTLQNMVLEYAHN 341
 QY 178 LIITNKLDGLKALKRVARTNGIK-----NAEETLINEVVRSTMQEELDAQTKT 227
 DB 342 LKVGEMLRVAVLAARQAGDAKWKQVLPILKASIKETLRLHPISVTLQ-----RY 393
 QY 228 TYVCDL-FRN 235
 DB 394 TVNDLVLRN 402

RESULT 7
 GTR2_RAT
 ID GTR2_RAT
 AC P12336;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Solute carrier family 2, facilitated glucose transporter, member 2
 DE (Glucose transporter type 2, liver).
 GN SLC2A2 OR GLUT2 OR GLUT-2.

DR	InterPro: IPR001179; FKBP_PPIase.
DR	Pfam: PF00036; ehband; 2.
DR	Pfam: PF00234; FKBP; 4.
DR	SMART: SM00054; EFh; 2.
DR	PROSITE: PS00453; FKBP_PPIASE_1; FALSE_NEG.
DR	PROSITE: PS00454; FKBP_PPIASE_2; 1.
DR	PROSITE: PS00059; FKBP_PPIASE_3; 4.
DR	PROSITE: PS00014; ER_TARGET; 1.
DR	PROSITE: PS00018; EF_HAND; 1.
KW	Isomerase; Retinase; Repeat; Glycoprotein; Signal; Phosphorylation;
KW	Endoplasmic reticulum; Calcium-binding.
FT	SIGNAL 1 33 POTENTIAL.
FT	CHAIN 34 581 65 KDA FK506-BINDING PROTEIN (POTENTIAL)
FT	DOMAIN 54 146 PPIASE, FKBP-TYPE 1.
FT	DOMAIN 179 258 PPIASE, FKBP-TYPE 1.
FT	DOMAIN 291 370 PPIASE, FKBP-TYPE 3.
FT	DOMAIN 389 482 PPIASE, FKBP-TYPE 4.
FT	CA_BIND 509 520 EF-HAND 1 (POTENTIAL).
FT	CA_BIND 554 565 EF-HAND 2 (POTENTIAL).
FT	CARBOHYD 69 69 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 351 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	STATE 578 581 PREVENT SECRETION FROM ER (POTENTIAL).
SO	SEQUENCE 581 AA; 64669 MW; 1B5IB3032089F555 CIRC64;

CC	-1	FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. PROBABLY A NEURONAL
CC		GLUCOSE TRANSPORTER.
CC	-1	SUBCELLULAR LOCATION: Integral membrane protein.
CC	-1	SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
CC		TRANSPORTER SUBFAMILY.
CC		-----
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration
CC		between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC		the European Bioinformatics Institute. There are no restrictions on its
CC		use by non-profit institutions as long as its content is in no way
CC		modified and this statement is not removed. Usage by and for commercial
CC		entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC		or send an email to license@isb-sib.ch).
CC		-----
DR	EMBL	M37785; AAA48662.1; -
DR	PIR	A41264; A41264.
DR	InterPro	IPR003663; Sugar_transporter.
DR	InterPro	IPR003662; sub_transporter.
DR	Pfam	PF00083; sugar_tr.1.
DR	PRINTS	PR00171; SUGRTRANSPORT.
DR	PROSITE	PS00216; SUGAR_TRANSPORT_1; 1.
DR	PROSITE	PS00217; SUGAR_TRANSPORT_2; 1.
KW		Transmembrane; Sugar transport; Transport; Glycoprotein;
KW		Multigene family.
FT	DOMAIN	1 27 6
FT	TRANSMEM	7 27
FT	DOMAIN	28 65
FT	TRANSMEM	66 86
FT	DOMAIN	87 94
FT	TRANSMEM	95 115
FT	DOMAIN	116 125
FT	TRANSMEM	126 146
FT	DOMAIN	147 154
FT	TRANSMEM	155 175
FT	DOMAIN	176 184
FT	TRANSMEM	185 205
FT	DOMAIN	206 271
FT	TRANSMEM	271 291
FT	DOMAIN	292 306
FT	TRANSMEM	307 327
FT	DOMAIN	328 335
FT	TRANSMEM	336 356
FT	DOMAIN	357 367
FT	TRANSMEM	368 388
FT	DOMAIN	389 398
FT	TRANSMEM	399 419
FT	DOMAIN	420 426
FT	TRANSMEM	427 447
FT	DOMAIN	448 496
FT	SITE	278 280
FT		
FT	CARBOHYD	44 44
SO	SEQUENCE	496 AA: 54174 MW: 753056PIVA79245 CRC64; N-LINKED (GLCNAC...)(POTENTIAL).

DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE DNA gyrase subunit B (EC 5.99.1.3) (Topoisomerase II subunit B).
 GN GYR B OR TOP2B OR TM0833.
 OS Thermotoga maritima.
 OC Bacteria; Thermotogales; Thermotoga.
 NCBI_TaxID=2336.
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB / DSM 3109;
 RX MEDLINE=97017137; PubMed=8863738;
 RA Guibaud O., Labedan B., Forterre P.;
 RT "A gyrB-like gene from the hyperthermophilic bacterium Thermotoga
 maritima";
 RL Gene 174:121-128(1996).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB / DSM 3109;
 RX MEDLINE=99287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
 Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
 Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 genome sequence of Thermotoga maritima";
 RL Nature 393:323-329(1999).
 CC -1- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-
 STRANDED DNA IN AN ATP-DEPENDENT MANNER AND ALSO CATALYZES THE
 INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED
 DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.
 CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
 of double-stranded DNA.
 CC -1- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA
 BREAKAGE AND REJOINING; THE B CHAIN CATALYZES ATP HYDROLYSIS. THE
 ENZYME FORMS AN A2B2 TETRAMER.
 CC -1- SIMILARITY: BELONGS TO THE TYPE II TOPOISOMERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U49692; AAC4498.1; -
 DR EMBL: AE001750; AAD35915.1; -
 DR HSSP: P06982; 1A16.
 DR TIGR: TM0833; -
 DR InterPro: IPR002288; DNA_gyraseB_C.
 DR InterPro: IPR001241; DNA_topoisomII.
 DR InterPro: IPR003594; HATPase_C.
 DR InterPro: IPR004359; HIS_KIN_sig.
 DR InterPro: IPR002936; TOPrim.
 DR Pfam: PF00986; DNA_gyraseB_C; 1.
 DR Pfam: PF00204; DNA_topoisomII; 1.
 DR Pfam: PF02518; HATPase_C; 1.
 DR Pfam: PF01751; TOPrim; 1.
 DR PRINTS: PRO00418; TP12FAMILY.
 DR ProDom: PD000616; DNA_topoisomII; 1.
 DR ProDom: PD149653; DNA_gyraseB_C; 1.
 DR SMART: SM00387; HATPase_C; 1.
 DR SMART: SM00433; TOP2C; 1.
 DR PROSITE: PS00177; TOPOISOMERASE_II; 1.
 DR PROSITE: PS00177; TOPOISOMERASE_II; 1.
 KM isomerase; Topoisomerase; ATP-binding; Complete proteome.
 FT CONFLICT 152 152 V -> M (in REF. 1).
 FT SEQUENCE 636 AA; 72427 MW; B21E0E0BC6F69E CRC64;

Query Match 6.0%; Score 84.5; DB 1; Length 636;
 Best Local Similarity 20.0%; Pred. No. 7;

Matches 57; Conservative 44; Mismatches 89; Indels 95; Gaps 13;
 QY 10 VGGGRFQMLVFLVPLPSIMLLIPHLLENPAALPGRCVHMLDNTGSGNETGLISE 69
 DB 26 IGSYGRKGLHLLV-----EVDNVSDEALAGCDWLRV-----TLHE 63
 QY 70 DALLRI-----SIPDSNLRPEKCRFVHPOMQHLNLG-----TI 105
 DB 64 DGSVEVEDNGRGIPIVD--IHEEGRSALEVFVTLHAGKRSKDSYKISGLHGVGSV 121
 QY 106 HSTSE-ADPEECVQGWVYDOSYFSTIWTKMDLVCDYSLKSVQF----- 150
 DB 122 NALSEMLEVRYNRDQKIRQKREKRPVTPVEVIGETKHGIVAFKDDPLFSETERDP 181
 QY 151 -----LLLTGLVGGI-----IGHVS-----DRMLVESAKWLLITNKDEGLAKLR 193
 DB 182 DLEHRLREIAFLVPLGLKIEFEDRINGEKTKFPFGGIVEYKRV-----INRGKALHD 235
 QY 194 V---ARTNGI--KNAEPLNIEV---VRSTWQEEELDAQOTKTV 229
 DB 236 VTIHKRTKVKTKNGEDEYVIELAFQYDYSIEDIYSPANTIKTV 280

RESULT 11
 ID GTR3_BOVIN STANDARD; PRT; 494 AA.
 AC P58352;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Solute carrier family 2, facilitated glucose transporter, member 3
 DE (Glucose transporter type 3, brain).
 GN SLC2A3 OR GLUT3.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 NCBI_TaxID=9913;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Augustin R., Fischer B.;
 RT "Glucose transporter expression during bovine preimplantation embryo
 development";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 RN (2)
 RP SEQUENCE OF 385-494 FROM N.A.
 RC TISSUE=Brain;
 RA Augustin R., Navarrete-Santos A., Fischer B.;
 RT "Expression of glucose transporters amongst ruminants";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. PROBABLY A NEURONAL
 GLUCOSE TRANSPORTER.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
 TRANSPORTER SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AY03938; AAK70222.1; -
 DR EMBL: AF308829; AAK63302.1; -
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
 DR Transmembrane; Sugar transporter; Transport; Glycoprotein;
 KW Multigene family.
 FT DOMAIN 1 5 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 6 26 1 (POTENTIAL).
 FT DOMAIN 27 64 EXTRACELLULAR (POTENTIAL).


```

FT TRANSMEM 65 85 2 (POTENTIAL).
FT DOMAIN 86 93 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 94 114 3 (POTENTIAL).
FT DOMAIN 115 124 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 125 145 4 (POTENTIAL).
FT DOMAIN 146 159 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 160 180 5 (POTENTIAL).
FT DOMAIN 181 183 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 184 204 6 (POTENTIAL).
FT DOMAIN 205 269 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 270 290 7 (POTENTIAL).
FT DOMAIN 291 305 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 306 326 8 (POTENTIAL).
FT DOMAIN 327 334 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 335 355 9 (POTENTIAL).
FT DOMAIN 356 363 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 364 384 10 (POTENTIAL).
FT DOMAIN 385 399 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 400 420 11 (POTENTIAL).
FT DOMAIN 421 424 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 425 445 12 (POTENTIAL).
FT DOMAIN 446 494 CYTOPLASMIC (POTENTIAL).
FT SITE 277 279 DEFINES SUBSTRATE SPECIFICITY (BY SIMILARITY).
FT CARBOHYD 43 43 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 358 358 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 385 389 PWETV -> SLVYC (IN REF. 2).
SQ SEQUENCE 494 AA: 34019 MW: EC42A3C648CAD23C CRC64;

```

```

Query Match 5.9%; Score 83.5; DB 1; Length 494;
Best Local Similarity 23.3%; Pred. No. 6.3;
Matches 27; Conservative 26; Mismatches 40; Indels 23; Gaps 3;

```

```

QY 135 WDLVCDYQSLKSVQVFLLTGMLVGGIIGHVSDRWLVESARWLITNKLDEGLK-ALRK 193
DB 184 WPLLGLFTLLPALIOCALP-----FCPESPRFLINRKKEEKAKELLOR 228
QY 194 VARTNGIKNAEETLINIVRSTMOEELDAQTKTYCDLFNPSMRKRICILVFLR 249
DB 229 LWGTEDVAODIOEKDESMRMSOEKOV-----TVLELFRAPNYPROPITISIMLQ 277

```

```

RESULT 12
GTR3_SHEEP STANDARD; PRT; 494 AA.
AC P47843;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Solute carrier family 2, facilitated glucose transporter, member 3
DE (Glucose transporter type 3, brain).
GN SLC2A3 OR GLUT3 OR GLUT-3.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COOPWORTH;
RX MEDLINE=96109471; PubMed=8653093;
RA Bennett B.L., Prosser C.G., Grigor M.R.;
RT "Isolation of cDNAs and tissue specific expression of ovine glucose
RT transporters.";
RL Biochem. Mol. Biol. Int. 37:9-16(1995).
RN [2]
RP SEQUENCE OF 88-248 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=97392487; PubMed=9250701;
RA Currie M.J., Bassett N.S., Gluckman P.D.;
RT "Ovine glucose transporter-1 and -3: cDNA partial sequences and
RT developmental gene expression in the placenta."

```

```

RL Placenta 18:393-401(1997).
CC -1- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. PROBABLY A NEURONAL
CC GLUCOSE TRANSPORTER.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN BRAIN.
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
CC TRANSPORTER SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L39214; AAC41629.1; -.
DR EMBL; U89030; AAB49313.1; -.
DR InterPro; IPR003663; Sugar_transporter.
DR InterPro; IPR003662; sub_transporter.
DR Pfam; PF00063; sugar_tr; 1.
DR PRINTS; PR00171; SUGRTNSPORT.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar transport; Transport; Glycoprotein;
KW Multigene family.
FT DOMAIN 1 5 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 6 26 1 (POTENTIAL).
FT DOMAIN 27 64 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 65 85 2 (POTENTIAL).
FT DOMAIN 86 93 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 94 114 3 (POTENTIAL).
FT DOMAIN 115 124 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 125 145 4 (POTENTIAL).
FT DOMAIN 146 153 5 (POTENTIAL).
FT TRANSMEM 154 174 6 (POTENTIAL).
FT DOMAIN 175 183 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 184 204 7 (POTENTIAL).
FT DOMAIN 205 269 8 (POTENTIAL).
FT TRANSMEM 270 290 9 (POTENTIAL).
FT DOMAIN 291 304 10 (POTENTIAL).
FT TRANSMEM 305 325 11 (POTENTIAL).
FT DOMAIN 326 334 12 (POTENTIAL).
FT TRANSMEM 335 355 1 (POTENTIAL).
FT DOMAIN 356 363 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 364 384 10 (POTENTIAL).
FT DOMAIN 385 399 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 400 420 11 (POTENTIAL).
FT DOMAIN 421 424 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 425 445 12 (POTENTIAL).
FT DOMAIN 446 494 CYTOPLASMIC (POTENTIAL).
FT SITE 277 279 DEFINES SUBSTRATE SPECIFICITY (BY SIMILARITY).
FT CARBOHYD 43 43 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 494 AA: 54194 MW: A89204D3EAV74BFBA CRC64;

```

```

Query Match 5.9%; Score 83.5; DB 1; Length 494;
Best Local Similarity 23.3%; Pred. No. 6.3;
Matches 27; Conservative 26; Mismatches 40; Indels 23; Gaps 3;

```

```

QY 135 WDLVCDYQSLKSVQVFLLTGMLVGGIIGHVSDRWLVESARWLITNKLDEGLK-ALRK 193
DB 184 WPLLGLFTLLPALIOCALP-----FCPESPRFLINRKKEEKAKELLOR 228
QY 194 VARTNGIKNAEETLINIVRSTMOEELDAQTKTYCDLFNPSMRKRICILVFLR 249
DB 229 LWGTEDVAODIOEKDESMRMSOEKOV-----TVLELFRAPNYPROPITISIMLQ 277

```

```

RESULT 13
C11A_CAPHI STANDARD; PRT; 520 AA.
ID C11A_CAPHI

```



```

FT TRANSMEM 429 449 12 (POTENTIAL).
FT DOMAIN 450 490 CYTOPLASMIC (POTENTIAL).
FT SITE 177 179 DEFINES SUBSTRATE SPECIFICITY (BY
FT CARBOHYD 44 44 SIMILARITY).
FT SEQUENCE 490 AA: 54086 MW: 638B176812D06104 CRC64;
SQ SEQUENCE 490 AA: 54086 MW: 638B176812D06104 CRC64;

Query Match
Best Local Similarity 23.3%; Score 82.5; DB 1; Length 490;
Matches 27; Conservative 28; Mismatches 38; Indels 23; Gaps 4;

QY 135 WLVCDYOSLKVYVDFLLITGMLVGGIIGHVSDFMLVESARMLIIT-NKLDEGIALRK 193
Db 185 WFLGGEFVFPALLOCIILP-----FAPESPRELLINRENNKAKSVLKK 229
QY 194 VARTNGIKNAEETLINEVVRSTMOEELDAAGTCTVCCDIFRNSMKRRIQIVFLR 249
Db 230 LRGITDVSSDLOEMK-EESRQMKREK-----KVTIMELFRSPMYRPIILAIVLQ 278

RESULT 15
GHT3_SCHPO STANDARD; PRI: 555 AA.
ID GHT3_SCHPO
AC 092339;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE High-affinity glucanase transporter ght3 (Hexose transporter 3).
GN GHT3 OR SPAC18.01.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
ON NCBI_TaxID=4896;
RX MEDLINE=20200352; PubMed=10735857;
RA Heland S., Radovanovic N., Hoefler M., Winderlickx J., Lichtenberg H.;
RT "Multiple hexose transporters of Schizosaccharomyces pombe.";
RL J. Bacteriol. 182:2153-2162(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: HIGH-AFFINITY GLUCONATE TRANSPORTER.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF051139; AAC63975.1; -
CC EMBL: 281312; CAB03595.1; -
CC InterPro: IPR003663; Sugar_transporter.
CC InterPro: IPR003662; sub_transporter.
CC Pfam: PF00063; sugar_tr; 1.
CC PRINTS: PR00171; SUGRTNSPORT.
CC PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
CC PROSITE: PS00217; SUGAR_TRANSPORT_2; 1.
CC Repeat: Transmembrane; Sugar transporter; Glycoprotein.
FT DOMAIN 1 9 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 10 30 1 (POTENTIAL).
FT DOMAIN 31 58 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 59 79 2 (POTENTIAL).
FT DOMAIN 80 87 CYTOPLASMIC (POTENTIAL).

```

```

FT TRANSMEM 88 108 3 (POTENTIAL).
FT DOMAIN 109 112 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 113 133 4 (POTENTIAL).
FT DOMAIN 134 144 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 145 165 5 (POTENTIAL).
FT DOMAIN 166 179 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 180 200 6 (POTENTIAL).
FT DOMAIN 201 266 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 267 285 7 (POTENTIAL).
FT DOMAIN 286 301 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 302 322 8 (POTENTIAL).
FT DOMAIN 323 328 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 329 349 9 (POTENTIAL).
FT DOMAIN 350 363 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 364 384 10 (POTENTIAL).
FT DOMAIN 385 404 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 405 425 11 (POTENTIAL).
FT DOMAIN 426 432 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 433 453 12 (POTENTIAL).
FT DOMAIN 454 555 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 523 555 SER-RICH.
FT SEQUENCE 555 AA: 62094 MW: 30DF04294D318DB CRC64;
SQ SEQUENCE 555 AA: 62094 MW: 30DF04294D318DB CRC64;

Query Match
Best Local Similarity 5.9%; Score 82.5; DB 1; Length 555;
Matches 52; Conservative 39; Mismatches 91; Indels 51; Gaps 11;

QY 22 VFILPSMLLIPHILLENFPAAPGRCWVHMLDNNSTSGNETGISDALRISIPIDS 81
Db 95 VYIYAEILLV-----TAVPS--WIOVLGKILAGVIGALS-----LSPGYOS 136
QY 82 NLREPKCRREFYHPQWQLHLNGTISTSEADTEPCVDGWDQSTFPSTIYTKMDLVCDY 141
Db 137 EVAPDQIRGAVVATYQ-----IFSTGAALVAACINNGTHTLRKTAS----NRTSFGI 184
QY 142 QSLKSVYQFLLITGMLVGGIIGHVSDFMLVESARMLIITNKLDEGIALRKVARFNGIK 201
Db 185 NMLMGI---LLMVGVL-----FLPESPRYLLYKGRDEALRLMCMMAELSPSS 229
QY 202 NAEETLINEVVRSTMOEELDAAGTCTVCCDIFRNSMKRRIQIVFLRKRI 252
Db 230 EIIOT-NFNTIKSDI--EIEWAGKARWIEIF-GMDIRYRCLGFLVLMFREL 278

Search completed: July 1, 2002, 11:43:04
Job time: 214 sec

```


RESULT 2

Query Match
 JMO089
 Organic cation transporter protein 2 - human
 N:Alternate names: OCTN2
 C:Species: Homo sapiens (man)
 C:Date: 18-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 28-May-1999
 C:Accession: JMO089
 R/Mu, X.; Prasad, P.D.; Leibaich, F.H.; Ganapathy, V.
 Biochem. Biophys. Res. Commun. 246: 589-595, 1998
 A:Title: CDNA sequence, transport function, and genomic organization of human OCTN2, a
 A:Reference number: JMO089; MUID:98289374
 A:Molecule type: mRNA
 A:Residues: 1-557 <MUA>
 A:Cross-references: GB:AF057164; NID:93273740; PIDN:ACC24828.1; PID:93273741
 A:Experimental source: Placenta
 C:Comment: This transporter functions in the elimination of cationic drugs and other xer

Query Match
 Best Local Similarity 13.8%; Score 193.5; DB 2; Length 557;
 Matches 83; Conservative 37; Mismatches 106; Indels 151; Gaps 12;

QY 3 FEELISQVGLGFRQMLHVFILPSMLILPH---ILENFAAIPGRHCVHMLDN-NT 58
 Db 4 YDEVTAFLDEMGPFQ--RLIFFLSAS-IIPNGFTGLSVFLNTPHRCRVPDANLSS 60
 QY 59 GSGNETGILSEDAALRISIPL---DSNLRPEKCRF---VHPQWOLILHN-GTHISTSEA 111
 Db 61 AMRNT-----VPLRLDRGREVPHSCRRYRATIANFSAIIGLEPRDVLQGL 108
 QY 112 DTEPCVDGWYDOSYFSTIVTKMDLYCDYOSLSKVQVFLTLGLVGIIGHSVSDR-- 169
 Db 109 EOECLDGMESQDVYISTVITVEMNLVCDMDKAPLTLISLFEVGLLSFISQGLSDRF 168
 QY 170 ----- 169
 Db 169 RKVNLVYMGMTQFSFLQIFSKNFMFVFLVGMQGISNYAARVLTGELIGKSVRI 228
 QY 170 ----- 169
 Db 229 ISTLGVCIFFAFYGVNLPFAIYIRDMRLVALTWPGVLCAVLMWFPEPSRMLISOG 288
 QY 183 RLDEGLKALRKVARNGKNAEETLNIEVRS---MOELDAQAQTKTVCDLFPNP 236
 Db 289 REEAEVYIRKAKANGI-----VVPSTIRPSELQDLSSKKQOSHNIIDLRTW 338
 QY 237 SMRRICILVFLKKKTS 253
 Db 339 NIRMVYIMIMMTIS 355

RESULT 3
 JEO346
 high-affinity carnitine transporter, CT1 - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
 C:Accession: JEO346
 R/Sekine, T.; Kusuhara, H.; Utsunomiya-Tate, N.; Tsuda, M.; Sugiyama, Y.; Kanai, Y.; Endo
 Biochem. Biophys. Res. Commun. 251: 586-591, 1998
 A:Title: Molecular cloning and characterization of high-affinity carnitine transporter
 A:Reference number: JEO346; MUID:99011422
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-557 <SEK>
 A:Cross-references: DDBJ:AB017260; NID:93869208; PIDN:BAA34399.1; PID:93869209

Query Match
 Best Local Similarity 13.7%; Score 193; DB 2; Length 557;
 Matches 83; Conservative 37; Mismatches 106; Indels 151; Gaps 12;

Query Match
 JMO089
 Organic cation transporter protein OCT1 - rat
 N:Alternate names: OCTN2
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 05-Nov-1999
 C:Accession: S50862; S78533; I58089
 R:Gordon, D.; Gordon, V.; Gambaryan, S.; Veyhl, M.; Koepsell, H.
 Nature 372: 549-552, 1994
 A:Title: Drug excretion mediated by a new prototype of polyspecific transporter.
 A:Reference number: I58089; MUID:95082907
 A:Accession: S50862
 A:Molecule type: mRNA
 A:Residues: 1-556 <GRU>
 A:Cross-references: EMBL:X78855; NID:9633621; PIDN:CAA5411.1; PID:9633622
 A:Experimental source: Kidney
 R:Gordon, V.G
 Submitted to the EMBL Data Library, January 1995
 A:Reference number: S78533
 A:Accession: S78533
 A:Molecule type: mRNA
 A:Residues: 1-542, 'N', 344-556 <GOR>
 A:Cross-references: EMBL:X78855; NID:9633621; PIDN:CAA5411.1; PID:9633622
 C:Keywords: glycoprotein; phosphoprotein; transmembrane protein
 F:20-44/Domain: transmembrane #status predicted <TM1>
 F:154-171/Domain: transmembrane #status predicted <TM2>
 F:178-197/Domain: transmembrane #status predicted <TM3>
 F:243-260/Domain: transmembrane #status predicted <TM4>
 F:267-283/Domain: transmembrane #status predicted <TM5>
 F:300-366/Domain: transmembrane #status predicted <TM6>
 F:380-398/Domain: transmembrane #status predicted <TM7>
 F:406-425/Domain: transmembrane #status predicted <TM8>
 F:435-452/Domain: transmembrane #status predicted <TM9>
 F:469-485/Domain: transmembrane #status predicted <TM10>
 F:494-514/Domain: transmembrane #status predicted <TM11>
 F:71.97/113,432/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:266.292/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status
 F:266.343,550/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status

Query Match
 Best Local Similarity 12.1%; Score 169.5; DB 2; Length 556;
 Matches 77; Conservative 47; Mismatches 95; Indels 157; Gaps 15;

QY 4 EELLISQVGLGFRQMLHVFILPSMLILPH---ILENFAAIPGRHCVHMLDN-NT 58
 Db 5 YDEVTAFLDEMGPFQ--RLIFFLSAS-IIPNGFTGLSVFLNTPHRCRVPDANLSS 60
 QY 59 GSGNETGILSEDAALRISIPL---DSNLRPEKCRF---VHPQWOLILHN-GTHISTSEA 111
 Db 61 AMRNT-----VPLRLDRGREVPHSCRRYRATIANFSAIIGLEPRDVLQGL 108
 QY 112 DTEPCVDGWYDOSYFSTIVTKMDLYCDYOSLSKVQVFLTLGLVGIIGHSVSDR-- 169
 Db 109 EOECLDGMESQDVYISTVITVEMNLVCDMDKAPLTLISLFEVGLLSFISQGLSDRF 168
 QY 170 ----- 169
 Db 169 RKVNLVYMGMTQFSFLQIFSKNFMFVFLVGMQGISNYAARVLTGELIGKSVRI 228
 QY 170 ----- 169
 Db 229 ISTLGVCIFFAFYGVNLPFAIYIRDMRLVALTWPGVLCAVLMWFPEPSRMLISOG 288
 QY 183 RLDEGLKALRKVARNGKNAEETLNIEVRS---MOELDAQAQTKTVCDLFPNP 236
 Db 289 REEAEVYIRKAKANGI-----VVPSTIRPSELQDLSSKKQOSHNIIDLRTW 338
 QY 237 SMRRICILVFLKKKTS 253
 Db 339 NIRMVYIMIMMTIS 355

```

OY 163 GGHVSDR----- 169
      |||
Db 167 VGYIADRGKRLCLLVTLTVSVSGVLTAAPDYTMLLFLLGMSKSGVSGYTLIT 226
OY 170 -----MLV-ES 174
      |||
Db 227 EFVSGYRRTAIIYQMAFYGVGLAGVAAVAPDMRMQLAVSLPTFLFLLYMFPES 286
OY 175 ARMLITNKLEDEGLKALRKAVATNG-INKAEETLIEVVRSTMOEBLDAQTKT-TVCDL 232
      |||
Db 287 PWWLSQKRTFAVIMQIAOKNGKVPAD-----LKMCLKEE-DASEKRSBSFADL 338
OY 233 FRNPSMRKRICILVF 248
      |||
Db 339 FRPTLRKHTVILMYL 354

RESULT 5
JC4884
Organic cation transporter protein 2 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1996 #sequence_revision 18-Oct-1996 #text_change 05-Nov-1999
C:Accession: J04884
R:Okuda, M.; Saito, H.; Urakami, Y.; Takano, M.; Inui, K.
Biochem. Biophys. Res. Commun. 224, 500-507, 1996
A:Title: cDNA cloning and functional expression of a novel rat kidney organic tra
A:Reference number: J04884; MUID:96295517
A:Accession: J04884
A:Molecule type: mRNA
A:Residues: 1-593 <OK>
A:Cross-references: DDBJ:D83044; NID:g1502282; PIDN:BA011754.1; PID:d1012421; PID:g15022
A:Experimental source: kidney
C:Comment: This protein is responsible for the transport of cationic drugs in kidney.

Query Match 11.6%; Score 162.5; DB 2; Length 593;
Best Local Similarity 18.3%; Pred. No. 2.1e-06;
Matches 68; Conservative 52; Mismatches 101; Indels 151; Gaps 12;

OY 4 EELLVGVGGIRGQMLHVFILPISLM-LLIPHLLENFAAIPGRHCW----- 50
      |||
Db 5 DDLEHIGEPHFQ-KQFFELALSGAFTPIYGVIGFTPDHHCSPGAATLSQRC 62
OY 51 -----VHMLDNTSGNETGILSEDALELRISIPDSNLRPKCRRTVHPOMQLHLNG 103
      |||
Db 63 GMSQAEELNVTYVGLSPDSASFSLQ--CMRYEV--DMNSTLDC--VDP-----LSS 109
OY 104 THISTSEADTEPCVDGWYDOSYFPSTIYTKMLVCDYOSLKVQVOLLTLGMLVGIIG 163
      |||
Db 110 LAADRQQLPLGPEHGMYNTP--GSSIVTEFNLCAHSWMLDLFOSVAVVGFPIGAMMI 167
OY 164 GHVSDR-----MLV----- 172
      |||
Db 168 GYLADRFGRKFLVTLILNAISGALMAISPNYAMMLVFRFGLGVSKAGMLIGYLITE 227
OY 173 -----ESA 175
      |||
Db 228 FVGLGYRRMVGICYQIAFTVGLLILAGVAVVPIPMRWLQFAVTLPNCFLLYFWCIPESP 287
OY 176 RMLITNKLEDEGLKALRKAVATNG-INKAEETLIEVVRSTMOEBLDAQTKT-TCDFRN 235
      |||
Db 288 RMLISQKRTFAVIMQIAOKNGKVPAD-----KSPVSLONLTPDEDA-GKKLKPSTLIDLVRT 341
OY 236 PSMRKRICILVF 247
      |||
Db 342 PDIRKHTVILMY 353

RESULT 6
T27870
Hypothetical protein ZK455.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans

```

```

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T27870
R:White, S.
submitted to the EMBL Data Library, November 1995
A:Reference number: Z20432
A:Accession: T27870
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-794 <MID>
A:Cross-references: EMBL:Z66567; PIDN:CA091492.2; GSPDB:GN00028; CESP:ZK455.8
A:Experimental source: clone ZK455
C:Genetics:
A:Gene: CESP:ZK455.8
A:Map position: X
A:Introns: 52/3; 119/2; 158/1; 207/3; 291/3; 335/2; 428/3; 610/2; 636/3; 733/3

Query Match 11.1%; Score 156; DB 2; Length 794;
Best Local Similarity 20.4%; Pred. No. 1.1e-05;
Matches 77; Conservative 46; Mismatches 103; Indels 152; Gaps 15;

OY 3 FEE-----LTSQVGLGRFQMLHVFILPISMLIPHLLENFAAAT---PGHRCVHML 54
      |||
Db 156 FEEDDGLIQLIGGCSYQI--IVYLISVQ-QVPHAMF-NLSVYVMYQPDHMKIPFF 211
OY 55 DNTSGNETGILSE--DALLR--ISIPDSN-----LREKC---RRVH-----PQWQ 97
      |||
Db 212 NERSFSA-ELGYNTYWDVQDLNSTIAFPRTFNKORNELHNDQCHYERDVHKLISWAO 270
OY 98 LHLN--GTIHSSEADTEPCVDGWYDOSYFPSTIYTKMLVCDYOSLKVQVOLLTLG 155
      |||
Db 271 VKQMNATGKINRKE-----WEIDTSVMRTIYIENNRVCDNMWSAHVHMSYSLG 321
OY 156 MLVGIIGGHVSDRW----- 170
      |||
Db 322 YLVGCFVYGFISDRYGRKATIGFGLITMLFGFLTYSKFEIFLYVRFLLAATNEAAL 381
OY 171 ----- 170
      |||
Db 382 AAYVLCMEVTGYRSTIVSGLIQAPWAGYAFALALVLRKSWTMHILCVLHIIISLM 441
OY 171 ---LVESARMLITNKLEDEGLKALRKAVATNGIKNAEETLIEVVRSTMOEBLDAQTKT 227
      |||
Db 442 LYFLPESPRLIILNKKTKQAEKIRIACHYN--KSLRPSDLGLVRAEKKKWKHNEKP 498
OY 228 TVCDLFRNPSMRKRICIL 245
      |||
Db 499 SYFHLFRSSSELFRFNVVL 516

RESULT 7
T30895
sugar transport protein homolog T08B1.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 29-Oct-1999
C:Accession: T30895
R:Henkhaus, J.; Wohlmann, P.; Leimbach, D.
submitted to the EMBL Data Library, August 1999
A:Description: The sequence of C. elegans cosmid T08B1.
A:Reference number: Z20926
A:Accession: T30895
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-591 <HEN>
A:Cross-references: EMBL:AF039039; PIDN:AAB94177.2
A:Experimental source: strain Bristol N2
C:Genetics:
A:Map position: V
A:Introns: 12/2; 62/1; 89/1; 111/3; 145/3; 325/2; 359/1; 468/2; 540/2
A:Note: T08B1.1

Query Match 7.7%; Score 108; DB 2; Length 591;

```

Best Local Similarity 19.8%; Pred. No. 0.14;
Matches 52; Conservative 45; Mismatches 97; Indels 68; Gaps 8;

QY 3 FEELLSQVGLGRFQMLHLVFLPDLMLIPHLLENFAAI--PGHRCVHMD----- 55
Db 32 FEELVQKIGAFGPYQIFCRIVIL-----YASIEWAGNSTFMHLIGSPED 76
QY 56 -NNTSGNETGIL--SEFALLRISIPDSNLRPEKCRFRVHPQWOLHLNGTHSTSEA 111
Db 77 WNCITLANNQTVIITAPHTDTCFICQNCNTLAPV----- 112
QY 112 DTEPCVDGKVVQOSTFPTITVKMDLVCDYQSLKSVQVFLTLGMLVGGIIGHVSDRW- 170
Db 113 -----QMLIEFF--SIVGQFOLICDDSDKVEYIEVIMAGSSILGSIIGMHGDFG 160
QY 171 -----LVESARWLTITNKLDGKALKRKVARTNGIKNAEELTIEVYRSTVOEELDAQTK 226
Db 161 ROTIFTEBELITITSMCTAAQSWIAFSVIGCV-NCFLYGVIEETSTLWMEFTSNKFR 219
QY 227 TTVCDLFRNPSMRKRICILVFL 248
Db 220 VIMVNAFQMPIAYMTALITLAWL 241

RESULT 8

C96758
Probable protein transporter T18K17.11 [Imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: C96758
R:Phenology: A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, J.; Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Hultz, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luo, J.S.; Maitl, R.; Marziani, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wang, W.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: C96758
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-539 <STO>
A:Cross-references: GB:AE005173; NID:965398860; PIDN:AAF18714.1; GSPDB:GNO041
C:Genetics:
A:Gene: T18K17.11
A:Map position: 1

Query Match 7.3%; Score 106; DB 2; Length 539;
Best Local Similarity 15.4%; Pred. No. 0.19;
Matches 48; Conservative 35; Mismatches 64; Indels 164; Gaps 7;

QY 1 MAREELLISQ-VGGIGRRQMLHVLIPSLMLIPHLLENFAAI--PGHRCVHMDNNTG 59
Db 43 LTVDEYIEQIHGAFOAILHALVIAWIFDAQTTLSIPSDAOPARLLA----- 94
QY 60 SGNETGISDALRLRISIPDSNLRPEKCRFRVHPQWOLHLNGTHSTSEADPEPCVDG 119
Db 95 -----TGAIIVGASL-----C-GLASGEMF----- 113
QY 120 WYVDQSTFPTITVKMDLVCDYQSLKSVQVFLTLGMLVGGIIGHVSDRWL----- 171
Db 114 WIGPKS---DTVSEWMLIQHKEFLVAVPSTLFFIGSLFGSGVGYLADSWFGKRTLL 170
QY 172 ----- 171
Db 171 SCVLTVTAFALFSPVWVYAFLEFRANGFRSGIGSCCIVLATEIYKRWGGOVGYGF 230
QY 172 -----VESARWLTITNKLDGSL 188

Db 231 FETVLGFLSLPMLAVLERKSRNLYRISFLPLGVAVCLLPFAESPRLWLVGRNKEAM 290

QY 189 KALKRVARTNG 199
Db 291 VLVKTKARLNG 301

RESULT 9

T16565
hypothetical protein K05F1.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16565
R:Wohlmann, P.
submitted to the EMBL data library, June 1995
A:Description: The sequence of C. elegans cosmid K05F1.
A:Reference number: Z18537
A:Accession: T16565
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-745 <WOH>
A:Cross-references: EMBL:U29377; NID:9868173; PID:9868176; PIDN:AAA68713.1; CESP:K05F
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:K05F1.6
A:Introns: 70/2; 160/3; 205/3; 249/3; 293/2; 350/1; 386/3; 462/3; 546/2; 669/3; 680/3

Query Match 7.3%; Score 103; DB 2; Length 745;
Best Local Similarity 18.2%; Pred. No. 0.53;
Matches 72; Conservative 44; Mismatches 112; Indels 168; Gaps 16;

QY 3 FEELLSQVGLGRFQMLHLVFLPDLMLIPHLLENFAAI--PGHRCVHMDNNTG 59
Db 113 FEGLINTIGGRWQI--WIVYLLAQ-QIPHAMENLWVYMYDPEFCWVPGF----- 164
QY 53 MLDNNTSGNETGILSE-----DALRISIP-----DSNLRPEKCRFRVHPQWOLHL 100
Db 165 ---NDT---NDTSSSGPYIMGVDDINKNISFVRPNANSDAQYQDSYFERSERRQ 218
QY 101 LN-----GTIHSTSEADPEP---CVDGMYVDQSTFPTITVKMDLVCDYQSLKSVQVFLTL 153
Db 219 LRRMPLETFAMSEAMKDVAPKRC--QAVHFEKDVWETIVDENLVCDSWFPAKGNHMYTS 277
QY 154 TGMVVGSIIGHVSDR----- 169
Db 278 TGYLGCVLGDIASDKIGRKPITIGFISLMSGLVFLPNDYPMFLIRLLSAICNEA 337
QY 170 -----W----- 170
Db 338 DLAAVTLQMEITGKYRAMVGSMLQAPWALGYALLALVITKSKRTIQVIAAGLHMSI 397
QY 171 -----LVESARWLTITNKLDGKALKRKVAR-----TNGIKNAEELTIEVYRST 215
Db 398 IFICSIPESPRWLMVONRSEAEVIRKACREPPFPMCTTSCGNLPSLELVSHRER 457
QY 216 MOEELDAQTKTVCDLFRNPSMRKR---ICILVFL 248
Db 458 KLNKKNNG--KIGFLDLFTYKELRYRTISVCI-VFM 490

RESULT 10

A31318
glucose transporter-like protein - human
C:Species: Homo sapiens (man)
C>Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 24-Sep-1999
C:Accession: A31318
R:Fukumoto, H.; Seino, S.; Imura, H.; Seino, Y.; Eddy, R.L.; Fukushima, Y.; Byers, M.
Proc. Natl. Acad. Sci. U.S.A. 85, 5434-5438, 1988
A:Title: Sequence, tissue distribution, and chromosomal localization of mRNA encoding
A:Reference number: A31318; MUID:88289735
A:Accession: A31318
A:Molecule type: mRNA

RESULT 12
F69587

RESULT 13

glucose transporter protein, hepatic - mouse
S06920
N:Alternate names: GLUT 2
C:Species: Mus musculus (house mouse)
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 24-Sep-1999
C:Accession: S06920; S05319; B4487; I48367
R:Stuzue, K.; Lodish, H.F.; Thorens, B.
Nucleic Acids Res. 17, 10099, 1989
A:Title: Sequence of the mouse liver glucose transporter.
A:Reference number: S06920; MUID:90098776
A:Accession: S06920
A:Molecule type: mRNA
A:Residues: 1-523 <SUZ>
A:Cross-references: EMBL:X16986; NID:955511; PIDN:CAA334855.1; PID:955512
R:Asano, T.; Shibasaki, Y.; Lih, J.L.; Ahanuma, Y.; Takaku, F.; Oka, Y.
Nucleic Acids Res. 17, 6386, 1989
A:Title: The nucleotide sequence of cDNA for a mouse liver-type glucose transporter p
A:Reference number: S05319; MUID:8936666
A:Accession: S05319
A:Molecule type: mRNA
A:Residues: 1-105; 'D', 107-180, 'T', 182-326, 'T', 328-350, 'F', 352-431, 'P', 433-523 <ASA>
R:Hogan, A.; Heyner, S.; Charron, M.J.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.;
Development 113, 363-372, 1991

A:Title: Glucose transporter gene expression in early mouse embryos.
 A:Reference number: A44887; MUID:92111400
 A:Accession: B44887
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 384-431; P: 433-496 <HOG>
 A:Cross-references: GB:S77926; NID:9242129; PIDN:AAB20847.1; PID:9242130
 A:Experimental source: embryo
 A>Note: sequence extracted from NCBI backbone (NCBI:77926, NCBI:77927)
 R:Wadner, G.; Thompson, N.; Haefliger, J.A.; Nicod, P.
 J. Biol. Chem. 269, 26912-26919, 1994
 A:Title: Characterization of the murine high Km glucose transporter GLUT2 gene and its t
 A:Reference number: A55078; MUID:95014557
 A:Accession: I48367
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-5 <RES>
 A:Cross-references: EMBL:X78722; NID:9517486; PIDN:CAA55368.1; PID:9683589
 C:Genetics:
 A:Gene: GLUT2
 C:Superfamily: glucose transport protein
 C:Keywords: liver; transmembrane protein

Query Match 6.88; Score 95; DB 2; Length 523;
 Best Local Similarity 25.88; Pred. No. 1.7;
 Matches 34; Conservative 27; Mismatches 36; Indels 36; Gaps 6;

OY 143 SIKSVQFLITGMLVGIIG-----GHVSDRWLV-----ESARW 177
 DB 186 ALGTHQALVYTGILISQIAGLSFILGN-QDWHMLLGLSAVPALLQCLLLFCESPRY 244
 OY 178 LIITNKLDGKALRKVARTNGINKNAEETLIEVVRSTMOEELDAQF--KTYCDLFRN 235
 DB 245 LYL--KLEEEVRAKSKLRIGTEDIYTKDIN-----EMRKEKEASTEQKVSIVQLFTD 296
 OY 236 PSMKRRCIIVFL 248
 DB 297 ANVRQPIVVALML 309

RESULT 14
 G84864
 probable membrane transporter [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
 C:Accession: G84864
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.;
 euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487
 A:Accession: G84864
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-521 <STO>
 A:Cross-references: GB:AE002093; NID:92289003; PIDN:AAB64332.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g43330
 A:Map position: 2
 C:Superfamily: glucose transport protein

Query Match 6.68; Score 93; DB 2; Length 521;
 Best Local Similarity 26.98; Pred. No. 2.6;
 Matches 35; Conservative 23; Mismatches 38; Indels 34; Gaps 7;

OY 125 SYFSTIVTK---MDLVCDYQSLKSVQFLITGMLVGIIGVSDRWLVESARWLI 180
 DB 188 SYLVNSAFYQVGTWRMGLGVSGVPAVIOFILML-----FMPESPRLWLFM 232
 181 TTKLDEGLKALKRVARTNGINKNAEETLIEVVRSTMOEELDAQOTKTV--CDLFRNPSM 238

DB 233 KNRKRAIQLV---ARTYDISRLD--EIDHLSAAEEEEE--KORRRTGYLDVFRSKEL 284
 OY 239 KRIICILVFL 248
 DB 285 R-----LAFIL 289

RESULT 15
 A31556
 glucose transport protein, hepatic - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 17-Nov-2000
 C:Accession: A31556; S68362
 R:Thorens, B.; Sarkar, H.K.; Kaback, H.R.; Lodish, H.F.
 Cell 55, 281-290, 1988
 A:Title: Cloning and functional expression in bacteria of a novel glucose transporter
 A:Reference number: A31556; MUID:89003066
 A:Accession: A31556
 A:Molecule type: mRNA
 A:Residues: 1-522 <THO>
 A:Cross-references: GB:J03145; NID:9204534; PIDN:AAA41298.1; PID:9204535
 R:Ahm, Y.; Kim, J.; Han, G.; Lee, B.; Kim, Y.
 Arch. Biochem. Biophys. 323, 387-396, 1995
 A:Title: Cloning and characterization of rat pancreatic beta-cell/liver type glucose
 A:Reference number: S68362; MUID:96063615
 A:Accession: S68362
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-246; K: 248-522 <AHN>
 A:Cross-references: EMBL:L28134
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, September 199
 C:Genetics:
 A:Insertions: 5/3; 36/3; 122/2; 164/1; 202/3; 257/1; 319/3; 354/3; 388/3; 456/3
 C:Superfamily: glucose transport protein
 C:Keywords: liver; transmembrane protein

Query Match 6.68; Score 93; DB 2; Length 522;
 Best Local Similarity 24.88; Pred. No. 2.6;
 Matches 33; Conservative 28; Mismatches 36; Indels 36; Gaps 6;

OY 143 SIKSVQFLITGMLVGIIG-----GHVSDRWLV-----ESARW 177
 DB 185 ALGTHQALVYTGILISQIAGLSFILGN-QDWHMLLGLSAVPALLQCLLLFCESPRY 243
 OY 178 LIITNKLDGKALRKVARTNGINKNAEETLIEVVRSTMOEELDAQF--KTYCDLFRN 235
 DB 244 LYL--NLEEEVRAKSKLRIGTEDIYTKDIN-----EMRKEKEASTEQKVSIVQLFTD 295
 OY 236 PSMKRRCIIVFL 248
 DB 296 PNVROPYVVALML 308

Search completed: July 1, 2002, 11:41:47
 Job time: 248 sec

PA (PROT-) PROTEGENE INC.
 XX Kato S, Kimura T;
 XX WPI: 2000-023358/02.
 DR P-PSDB; AAY52386.
 XX Human proteins with transmembrane domains, involved in control of cell
 PT proliferation and differentiation, useful for treating e.g. cancer or
 PT inflammation
 PS Claim 3; Page 84; 114pp; English.
 XX This sequence represents the coding sequence of human CDNA clone
 CC HP02000 which encodes a 32 kD protein with two putative transmembrane
 CC domains. The CDNA was isolated from a human liver cell CDNA library, and
 CC from tissue localisation studies has been found to be expressed only in
 CC the liver. The protein has homology with the rat organic cation
 CC transporter (EMBL Accession No. Y09945) which is involved in drug
 CC excretion, and may have a similar function. The protein
 CC may be used to raise specific antibodies, as assay reagents, as
 CC diagnostic tissue markers, for the isolation of cognate receptors,
 CC ligands and binding proteins, and as biologically active agents.
 CC Nucleotides encoding the protein may be used as primers and probes or
 CC antisense molecules, and in gene therapy. Cells transformed with these
 CC nucleotides may be used to screen for agonists and antagonists which are
 CC potentially useful therapeutically.
 XX Sequence 804 BP; 226 A; 171 C; 193 G; 214 T; 0 other;

Query Match 100.0%; Score 804; DB 21; Length 804;
 Best Local Similarity 100.0%; Pred. No. 1.9e-248;
 Matches 804; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggccttggagagctcttgatgaagttggagccttgggaatttcagatcttcac 60
 DB 1 atggccttggagagctcttgatgaagttggagccttgggaatttcagatcttcac 60
 QY 61 ctggttttattcttcctccctccatgtatataccctataactcgtagaacttt 120
 DB 61 ctggttttattcttcctccctccatgtatataccctataactcgtagaacttt 120
 QY 121 gctgcagcattcctcgtcgtcgttgcgtgcacatgctgacaaataactgacatc 180
 DB 121 gctgcagcattcctcgtcgtcgttgcgtgcacatgctgacaaataactgacatc 180
 QY 181 ggttaagaactcgtgaatcctcagtgagaatgcctcttgagaatctatccactagac 240
 DB 181 ggttaagaactcgtgaatcctcagtgagaatgcctcttgagaatctatccactagac 240
 QY 241 tcaaatctgagcagagaagtgctgcgttgcattgccacccagtgagcctcttcac 300
 DB 241 tcaaatctgagcagagaagtgctgcgttgcattgccacccagtgagcctcttcac 300
 QY 301 ctgaatcggagactacacagcacaagtgagcagacacacacctgtgtgagtctg 360
 DB 301 ctgaatcggagactacacagcacaagtgagcagacacacacctgtgtgagtctg 360
 QY 361 ggtatgatacaagctactcctcgtcagcattgtgactaagtgagcctggtatgac 420
 DB 361 ggtatgatacaagctactcctcgtcagcattgtgactaagtgagcctggtatgac 420
 QY 421 tatagctcctgaatcagtggttcaattcctcctcgtcagcagtgagtgaggagc 480
 DB 421 tatagctcctgaatcagtggttcaattcctcctcgtcagcagtgagtgaggagc 480
 QY 481 accatagtgagcagatgtcacaagcagtggtgctggaactcgtcgtgttatatac 540
 DB 481 accatagtgagcagatgtcacaagcagtggtgctggaactcgtcgtgttatatac 540
 QY 541 accaataaactagatgaggtctaaagcacttagaaagtgtgacgcacaaatgagata 600
 DB 541 accaataaactagatgaggtctaaagcacttagaaagtgtgacgcacaaatgagata 600

DB 541 accaataaactagatgaggtctaaagcacttagaaagtgtgacgcacaaatgagata 600
 QY 601 aagaatgctgaagaacccttgacatagaagttgttaagatccaccttgacggagctg 660
 DB 601 aagaatgctgaagaacccttgacatagaagttgttaagatccaccttgacggagctg 660
 QY 661 gatgcagcacaagcacaactactcgtgtgactgttccgcacccagtatcgctaa 720
 DB 661 gatgcagcacaagcacaactactcgtgtgactgttccgcacccagtatcgctaa 720
 QY 721 aggatcgtatcctgtgatttttgagaaaaaatctcaaggaagagcataaatgat 780
 DB 721 aggatcgtatcctgtgatttttgagaaaaaatctcaaggaagagcataaatgat 780
 QY 781 tgcacacaaagtacccaattt 804
 DB 781 tgcacacaaagtacccaattt 804

RESULT 2
 ID AAZ38317 standard; CDNA; 1705 BP.
 XX AAZ38317;
 AC AAZ38317;
 XX 09-FEB-2000 (first entry)
 DE Human transmembrane protein CDNA clone HP02000.
 KW HP02000; transmembrane domain; liver; expression; homology;
 KW organic cation transporter; drug excretion; antibody; assay reagent;
 KW diagnostic marker; primer; probe; antisense; gene therapy;
 KW agonist; antagonist; ligand; therapeutic; ds.
 OS Homo sapiens.

XX Key Location/Qualifiers
 FH 187..993
 FT CDS /tag= a
 FT /product= "Human transmembrane protein HP02000"
 XX W09955862-A2.
 XX 04-NOV-1999.
 XX 27-APR-1999; 99WO-JP02226.
 XX 28-APR-1998; 98JP-0119395.
 XX (SAGA) SAGAMI CHEM RES CENT.
 PA (PROT-) PROTEGENE INC.
 XX Kato S, Kimura T;
 XX WPI: 2000-023358/02.
 DR P-PSDB; AAY52386.

PT Human proteins with transmembrane domains, involved in control of cell
 PT proliferation and differentiation, useful for treating e.g. cancer or
 PT inflammation
 PS Claim 4; Page 89-91; 114pp; English.

XX This sequence represents the human CDNA clone HP02000 which
 CC encodes a 32 kD protein with two putative transmembrane
 CC domains. The CDNA was isolated from a human liver cell CDNA library, and
 CC from tissue localisation studies has been found to be expressed only in
 CC the liver. The protein has homology with the rat organic cation
 CC transporter (EMBL Accession No. Y09945) which is involved in drug
 CC excretion, and may have a similar function. The protein
 CC may be used to raise specific antibodies, as assay reagents, as
 CC diagnostic tissue markers, for the isolation of cognate receptors,
 CC ligands and binding proteins, and as biologically active agents.

CC Nucleotides encoding the protein may be used as primers and probes or
CC antisense molecules, and in gene therapy. Cells transformed with these
CC nucleotides may be used to screen for agonists and antagonists which are
CC potentially useful therapeutically.

SQ Sequence 1705 BP; 559 A; 305 C; 344 G; 497 T; 0 other;

Query Match	100.0%	Score 804:	DB 21:	Length	1705:
Best Local Similarity	100.0%	Pred. No. 2	9e-248:		
Matches	804:	Conservative	0:	Mismatches	0:
				Indels	0:
				Gaps	0:

QY	1	atgacgtttaaagagacgtcttgaaatcaagtgtggagccttggagattcaagatgctaac	60
Db	187	atgacgtttaaagagacgtcttgaaatcaagtgtggagccttggagattcaagatgctaac	246
QY	61	ctggatttattctctccctctccatgltatataccctcatactactgtagaagactt	120
Db	247	ctggatttattctctccctctccatgltatataccctcatactactgtagaagactt	306
QY	121	gtcgagacattcctctgttcaatcggtgtggtcccaatgtgtgacaatatatactgatact	180
Db	307	gtcgagacattcctctgttcaatcggtgtggtcccaatgtgtgacaatatatactgatact	366
QY	181	ggtaatgaaactggaatccctccaatgaaagacccctctggaatctctatacccaactagac	240
Db	367	ggtaatgaaactggaatccctccaatgaaagacccctctggaatctctatacccaactagac	426
QY	241	tcaaactctgagccagagaaagtgcgtgcgttgtccatcccaatgtagcagcttcttcaac	300
Db	427	tcaaactctgagccagagaaagtgcgtgcgttgtccatcccaatgtagcagcttcttcaac	486
QY	301	ctggaatggagctatccacacacaaagtgaaaggagagacacaaacccgtgtgtagtgctgg	360
Db	487	ctggaatggagctatccacacacaaagtgaaaggagagacacaaacccgtgtgtagtgctgg	546
QY	361	gtatatatcaaaagcttactccctctcgacacattgtgtgactaagtgtggagccctgtgata	420
Db	547	gtatatatcaaaagcttactccctctcgacacattgtgtgactaagtgtggagccctgtgata	606
QY	421	tatcaatgccaatgaatacgaatggtgttcaattctctactcttgacctgtgaatgctgcgtgtggagac	480
Db	607	tatcaatgccaatgaatacgaatggtgttcaattctctactcttgacctgtgaatgctgcgtgtggagac	666
QY	481	atcatagtgtgcacatgctctcagaaaggtgctagtggaatcctctcgtgtgttgatatac	540
Db	667	atcatagtgtgcacatgctctcagaaaggtgctagtggaatcctctcgtgtgtgttgatatac	726
QY	541	accaataaactagatgtgaggggttaaaagcacttagaaaaagltgcagccacaaatgtgata	600
Db	727	accaataaactagatgtgaggggttaaaagcacttagaaaaagltgcagccacaaatgtgata	786
QY	601	aagaatgtctgaagaacaccctggaataatagaggtgtgtaagatcccaatgcagagggagctg	660
Db	787	aagaatgtctgaagaacaccctggaataatagaggtgtgtaagatcccaatgcagagggagctg	846
QY	661	gatcagacagacacaaacactacctgtgtgtgactgtgtccgcaaccccaatgtagttaa	720
Db	847	gatcagacagacacacaaacactacctgtgtgtgactgtgtccgcaaccccaatgtagttaa	906
QY	721	aggaatctgtatcctctgtaattttgagaaaaaaaatctcaaggaaagagcataaaaatgtat	780
Db	907	aggaatctgtatcctctgtaattttgagaaaaaaaatctcaaggaaagagcataaaaatgtat	966
QY	781	tgctacacaaaagtgacaaattt 804	
Db	967	tgctacacaaaagtgacaaattt 990	

RESULT	3
AAH98847	
ID	AAH98847 standard; cDNA; 1030 BP.
XX	

AC	AAH98847;	
XX		
DT	12-OCT-2001	(first entry)
XX		
DE	Human EST-derived coding sequence	SEQ ID NO: 704

KM Human; sheep; pig; cow; fruitfly; yeast; hamster; macaque; horse
KM tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KM diagnostics; forensic test; gene mapping; genetic disorder;
KM biodiversity; gene therapy; nutrition; ss.

OS Homo sapiens

PN WO200154477-A2.

PD 02-AUG-2001.

PF 25-JAN-2001; 2001WO-US02687

PR 25-JAN-2000; 2000US-0491404
17-TUE-2000; 2000US-0617745

PR 03-AUG-2000; 2000US-0631451
1E-SEP-2000; 2000TC-0653870

XX
XX

SECRET & UNCLASSIFIED

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

PI	Cao Y,	Drmanac R,
yy		

DR WPI; 2001-476164/51.
D-NDP; 28M34198

DT	XX	Testated polymerside
----	----	----------------------

PJ antipodes and rese

CLALM L; Page 030; -
PS
XX

proteins from a variety of present inventors

CC Cow, pig, mussels, 7

from the organism of
forensics, gene map

CC adversely and to
CC of the invention.

aa Sequence 1030 BP; 21

Best Local Similarity

[illegible][illegible]

22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051

Every Match	97.3%	Score 782;	DB 22;	Length 1030;
1st Local Similarity	-99.8%;	Pred. No. 2.7e-241;		
Matches 804;	Conservative 0;	Mismatches 0;	Indels 2;	Gaps 2.

QY	1	atggcctttgagagactctttaagcaagtcttgagagccttggagaattcagatgcgttcat	60
Db	186	atggcctcttgagagactcttgagtcgaagcttggagagcttggagattctgagatgcttcat	245
QY	61	ctggcttttatctctccctctctcaatgtaattatccctacatatactgtagagaacttt	120
Db	246	ctggcttttatctctccctctctcaatgtaattatccctacatatactgtagagaacttt	305
QY	121	gcttgagacattccctggtgcatacgtttgctgggttcacaaatgcctggagcaataaactgagct	180
Db	306	gcttgagacattccctggtgcatacgtttgctgggttcacaaatgcctggagcaataaactgagct	365
QY	181	ggttaatgaactcgtgatactctcagttgaagaatgcctcttgagaaatctctatccacatgac	240
Db	366	ggttaatgaactcgtgatactctcagttgaagaatgcctcttgagaaatctctatccacatgac	425
QY	241	tcaatctgagagccgagagaagatgctgtgcgttttgcaatccccagttggagactcttcac	300
Db	426	tcaatctgagagccgagagaagatgctgtgcgttttgcaatccccagttggagactcttcac	485

QY 301 ctgaatggagctatccacagcaaaagtgagcagacacagaccctgtgtgagtgctgg 360
 Db 486 ctgaatggagctatccacagcaaaagtgagcagacacagaccctgtgtgagtgctgg 545
 QY 361 gatatgtatcaaaagctacttcccttcgacacattgtgactaaagtggagcctgtatgtat 420
 Db 546 gatatgtatcaaaagctacttcccttcgacacattgtgactaaagtggagcctgtatgtat 605
 QY 421 tatcagctacgaatcaatggtgttcacatctcactctcactctgactgaatgctggggagggc 480
 Db 606 tatcagctacgaatcaatggtgttcacatctcactctcactctgactgaatgctggggagggc 665
 QY 481 atcataagtgagccatgctcagacacaggtgctggtggaatctcgtctggtgtgataatc 540
 Db 666 atcataagtgagccatgctcagacacaggtgctggtggaatctcgtctggtgtgataatc 725
 QY 541 accataaactagatgagggccttaagagcacttagaagaagttgcagcagacaatgtgata 600
 Db 726 accataaactagatgagggccttaagagcacttagaagaagttgcagcagacaatgtgata 785
 QY 601 aagaatgctg-aagaacccctgaacacatagaggttgaagatccacatcgagagagagct 659
 Db 786 aagaatgctgagaaagaaacccctgaacacatagaggttgaagatccacatcgagagagagct 845
 QY 660 ggatgagcagacagacacaaactactgtgtgtgactgtgtccgcgaacccagatgctgtaa 719
 Db 846 ggatgagcagacagacacaaactactgtgtgtgactgtgtccgcgaacccagatgctgtaa 905
 QY 720 aagatctgctatcctggtatcttggag-aaaaaaatcctcaaggaaagcataaaatg 778
 Db 906 aagatctgctatcctggtatcttggagaaaaaaatcctcaaggaaagcataaaatg 965
 QY 779 atgtctacacaaaagtgaccaaattt 804
 Db 966 atgtctacacaaaagtgaccaaattt 991

RESULT 4

ABA08976 standard; cDNA; 1353 BP.

XX ABA08976;

DT 11-JAN-2002 (first entry)

DE Human secreted protein homologue-encoding cDNA, SEQ ID NO:752.

XX Human: cytokine; cell proliferation; cell differentiation; growth factor;
 XX hematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytosatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnery; antilucer; ss.

OS Homo sapiens.

XX WO200157188-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001MO-US03800.

XX 03-FEB-2000; 2000US-0496914.

XX 27-APR-2000; 2000US-0560875.

XX (HYSF-) HYSFO INC.

XX Tang YT, Liu C, Dzumanac RT;
 XX WPI: 2001-457740/49.
 DR P-PSDB: ABB11732.
 DR
 XX
 PT Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 XX e.g. arthritis and cancer -
 XX
 PS Claim 1; Page 677; 1963pp; English.

Sequences ABB10981-ABBI2330 represent 1350 novel human polypeptides, and sequences ABA08225-ABA09574 represent nucleic acids encoding them. The invention also relates to vectors and recombinant host cells comprising a nucleotide of the invention, methods of producing the novel polypeptides, antibodies against the polypeptides, methods of detecting the nucleotides or polypeptides in a sample, and methods of identifying compounds which bind to polypeptides of the invention. Although novel, many of the polypeptides of the invention have homology to known proteins, thereby giving an insight into their probable biological activities, and hence potential therapeutic applications. The polypeptides of the invention may have various activities, including cytokine, cell proliferation or cell differentiation activities; stem cell growth factor activity; hematopoiesis regulatory activity; tissue growth activity; immunomodulatory activity; activin- or inhibin-related activities; chemotactic or chemokinetic activities; haemostatic, thrombotic or thrombolytic activities; receptor or ligand activities; or may be involved in oncogenesis, cancer cell proliferation or metastasis. Depending on their biological activities, polypeptides and nucleotides of the invention are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Such conditions include cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell disorders), chronic inflammatory conditions (e.g., asthma or arthritis), proliferative retinopathy, atherosclerosis, coronary heart disease, arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal vascular growth. Polypeptides involved with tissue regeneration and repair (or nucleic acids encoding them) may be used to promote wound healing (e.g., of burns, incisions and ulcers), while those with CC immunomodulatory activities may be used in the treatment of viral, bacterial and fungal infections in addition to immune disorders. CC Polypeptides with growth factor activity may be used in cell cultures to promote cell growth. For example, such polypeptides may be used to CC manipulate stem cells in culture to give rise to neuroepithelial cells CC that can be used to augment or replace cells damaged by illness. CC autoimmune disease or accidental damage. The polypeptides and nucleotides CC may also be used in the diagnosis of the above conditions, and in drug CC screening techniques. The present sequence represents a cDNA encoding a novel human polypeptide of the invention.

Sequence 1353 BP; 349 A; 298 C; 315 G; 391 T; 0 other;

Query Match 63.28; Score 508; DB 22; Length 1353;
 Best Local Similarity 100.0%; Pred. No. 6, 2e-153;
 Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggcctttagagagccttgatgacaaagtggagccttggagaaattcaatgctcat 60
 Db 186 atggcctttagagagccttgatgacaaagtggagccttggagaaattcaatgctcat 245
 QY 61 ctggttttattcttccctctcctatgtatlaatccctcatatcgtgagaaactt 120
 Db 246 ctggttttattcttccctctcctatgtatlaatccctcatatcgtgagaaactt 305
 QY 121 gctgcagcattcctcgtgcatcgttctggtgctcagatgctggaacaaataactgact 180
 Db 306 gctgcagcattcctcgtgcatcgttctggtgctcagatgctggaacaaataactgact 365
 QY 181 ggaatgaactggaatccctcagtgaaatgagcctcttggaactctatccactgac 240
 Db 366 ggaatgaactggaatccctcagtgaaatgagcctcttggaactctatccactgac 425

QY 241 tcaaatctggagccagaagatgctgcgtctgtccatcccaatggcagcttctac 300
 Db 426 tcaaatctggagccagaagatgctgcgtctgtccatcccaatggcagcttctac 485
 QY 301 ctgaatggagctatcccaagacagaagtgagcagacacagaacctgtgtgagctg 360
 Db 486 ctgaatggagctatcccaagacagaagtgagcagacacagaacctgtgtgagctg 545
 QY 361 gtatatgatacaagctactctccctcgacacattgtgactgaatggagcctgtatgat 420
 Db 546 gtatatgatacaagctactctccctcgacacattgtgactgaatggagcctgtatgat 605
 QY 421 tatcagtcacatgaatcagtggttcaatctcctactctgactgtaatgctggagagc 480
 Db 606 tatcagtcacatgaatcagtggttcaatctcctactctgactgtaatgctggagagc 665
 QY 481 atcatagtgccatgctcagacagct 508
 Db 666 atcatagtgccatgctcagacagct 693

RESULT 5

AAC85824 standard; DNA; 2684 BP.

AAC85824;

06-AUG-2001 (first entry)

hOAT5 DNA.

Human: organic anion transporter; hOAT; liver; kidney;

Membrane protein; transport; organic anion; splice variant; ds.

Homo sapiens.

Key Location/Qualifiers

CDS 190..1825

FT /*tag= a

FT /product= "hOAT5"

MO200104283-A2.

18-JAN-2001.

12-JUL-2000; 2000MO-US18980.

12-JUL-1999; 99US-0143771.

(META-) METABASIS THERAPEUTICS INC.

Sun W;

WPI: 2001-367057/38.

P-PEDB; AAB47276.

Nucleic acids encoding human organic anion transporter polypeptides,

useful in gene therapy procedures -

Claim 1; fig 6; 95pp: English.

The sequences given in AAC85819-24 encode human organic anion

transporter (hOAT) polypeptides. hOAT polypeptides are preferentially

expressed in the liver and kidneys of humans. OAT's are membrane

proteins that facilitate the transport of organic anions across the

cell membrane. The mechanism of transport is thought to be a secondary

active transport involving exchange of another organic anion.

hOAT2a and hOAT2b are thought to be splice variants as they are

identical except at the C-terminal end. hOAT proteins and the DNA

Sequence 2684 BP; 800 A; 533 C; 541 G; 810 T; 0 other;

Query Match 63.2%; Score 508; DB 22; Length 2684;
 Best Local Similarity 100.0%; Pred. No. 9 Le-153;
 Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atggcctttgagagcctctgagtcagatgtgagagccttggagatattagatgtctac 60
 Db 200 atggcctttgagagcctctgagtcagatgtgagagccttggagatattagatgtctac 259
 QY 61 ctggtttattcttccctctctatgtttatataccctaatatctgctagaacctt 120
 Db 260 ctggtttattcttccctctctatgtttatataccctaatatctgctagaacctt 319
 QY 121 gcttgaagccattccctgctatcgttgcctggttccacatgctgagcaataatagatct 180
 Db 320 gcttgaagccattccctgctatcgttgcctggttccacatgctgagcaataatagatct 379
 QY 181 ggtaatgaaactggaatctctcagatgaatgagatgagcctcttgagaatctatccactagac 240
 Db 380 ggtaatgaaactggaatctctcagatgaatgagatgagcctcttgagaatctatccactagac 439
 QY 241 tcaaatctggagccagaagatgctgcgtctgtccatcccaatggcagcttctac 300
 Db 440 tcaaatctggagccagaagatgctgcgtctgtccatcccaatggcagcttctac 499
 QY 301 ctgaatggagctatcccaagacagaagtgagcagacacagaacctgtgtgagctg 360
 Db 500 ctgaatggagctatcccaagacagaagtgagcagacacagaacctgtgtgagctg 559
 QY 361 gtatatgatacaagctactctccctcgacacattgtgactgaatggagcctgtatgat 420
 Db 560 gtatatgatacaagctactctccctcgacacattgtgactgaatggagcctgtatgat 619
 QY 421 tatcagtcacatgaatcagtggttcaatctcctactctgactgtaatgctggagagc 480
 Db 620 tatcagtcacatgaatcagtggttcaatctcctactctgactgtaatgctggagagc 679
 QY 481 atcatagtgccatgctcagacagct 508
 Db 680 atcatagtgccatgctcagacagct 707

RESULT 6

AA159215 standard; cDNA; 1375 BP.

AA159215;

22-OCT-2001 (first entry)

Human polynucleotide SEQ ID NO 1418.

Human: nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

peripheral nervous system; neuropathy; central nervous system; CNS;

Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;

amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

chemokine; thrombolytic; drug screening; arthritis; inflammation;

leukemia; ss.

Homo sapiens.

WO200153312-A1.

26-JUL-2001.

26-DEC-2000; 2000MO-US34263.

21-JAN-2000; 2000US-0488725.

23-APR-2000; 2000US-0552317.

09-JUL-2000; 2000US-0598042.

19-JUL-2000; 2000US-0620312.

03-AUG-2000; 2000US-0653450.

CC disease, stroke, cerebral neoplasms, Pick's disease, Huntington's
 CC disease and Parkinson's disease, demyelinating diseases, mental disorders
 CC including mood, anxiety, schizophrenia and seasonal affective disorder,
 CC muscle disorder including cardiomyopathy, myocarditis, polymyositis,
 CC dermatomyositis, arrhythmias and asthma and immunological disorders
 CC including AIDS, adult respiratory distress syndrome (ARDS), allergies,
 CC anaemia, diabetes mellitus, rheumatoid arthritis, scleroderma, Sjogren's
 CC syndrome, systemic lupus erythematosus and other diseases including
 CC sickle cell anaemia, Wilson's disease, cataracts, infertility, pulmonary
 CC artery stenosis, Grave's disease, Cushing's disease, Addison's disease,
 CC glucose-galactose malabsorption syndrome, hypercholesterolaemia, cancers
 CC psoriasis and viral, bacterial, fungal, helminthic and protozoal
 CC infections. TRICH DNA is useful in gene therapy and in diagnostic
 CC purposes.

Sequence 1986 BP; 505 A; 496 C; 457 G; 528 T; 0 other;

Query Match 41.7%; Score 335.2; DB 22; Length 1986;
 Best Local Similarity 78.7%; Pred. No. 3.7e-97;
 Matches 400; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 1 atggccttggagagcctcttgagcaagttggagccttggagagattcagatgcttcac 60
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 248 atggccttggagagcctcttgagcaagttggagccttggagagattcagatgcttcac 307
 QY 61 ctggtttttttcttcctcctcctcatgttatatccctcctatctcgtcagaaacttt 120
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 308 actgttttctcctcaatctctgtctgtgacatcctcttattatgtggaacttc 367
 QY 121 gctcagcattccctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 180
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 368 actgcatctactcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 427
 QY 181 ggtatgaactggaatcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 240
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 428 gaaatgacactggggcctcagcagcagatgacacttggaaatcctcctcctcctcctcct 487
 QY 241 tcaaatctggagcagagagatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 300
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 488 tcaaatctggagcagagagatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 547
 QY 301 ctgaatggagcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 360
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 548 ctgaatggagcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 607
 QY 361 gttatgataaagctactcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 420
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 608 gttatgataaagctactcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 667
 QY 421 tatcagtcactgaatgaatggttcaattctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 480
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 668 tctcaactcctgactctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 727
 QY 481 atcagatgtggcactgtctcagacagt 508
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 728 atcctaggcggtcattatcagacagt 755

RESULT 8
 AAC61892
 ID AAC61892 standard; cDNA; 2027 BP.

AC AAC61892;

XX 06-MAR-2001 (first entry)

XX cDNA encoding a human secreted protein.

KW Secreted protein; platelet disorder; stem cell disorder; osteoporosis;
 KW osteoarthritis; burn; incision; ulcer; periodontal disease; neuropathy;
 KW nervous system disease; bone growth; cosmetic plastic surgery;
 KW gut protection; gut regeneration; fibrosis; cancer;

KW bone marrow transplantation; ss.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT CDS 270..1289
 FT /tag= a
 FT /product= "secreted protein"

MO200061755-A2.

19-OCT-2000.

10-APR-2000; 2000MO-0509555.

09-APR-1999; 99US-0128574.

20-AUG-1999; 99US-0150054.

(CHIR) CHIRON CORP.
 Garcia PD;

WPI: 2000-665133/64.
 P-PSDB; AAB19406.

Novel secreted human proteins useful for stimulating blood cell
 generation in patients receiving cancer chemotherapy, treating bone
 marrow transplantation patients and for healing fractured bones -
 Claim 2; Page 71-72; 74pp: English.

AAC61879-93 encode secreted human proteins. The secreted proteins are
 useful in assays to determine their biological activities. The proteins
 can also be used as biomarkers to identify tissues or cell types which
 express the proteins. The polynucleotide molecules can be used as
 biomarkers for tissues or chromosomes and to elicit immune responses.
 The proteins and antibodies are useful in diagnosis and treatment of
 diseases associated with altered expression of these proteins. The
 proteins are also useful for prevention or treatment of platelet
 disorders, stem cell disorders, osteoporosis or osteoarthritis, burns,
 incisions, ulcers, periodontal diseases, central and peripheral nervous
 system diseases and neuropathies, for healing fractured bones and to
 induce cartilage and/or bone growth in cosmetic plastic surgery. The
 proteins are also useful for gut protection or regeneration, for the
 treatment of lung or liver fibrosis, for stimulating blood cell
 generation in patients receiving cancer chemotherapy and for treatment
 of bone marrow transplantation patients.

Sequence 2027 BP; 530 A; 501 C; 459 G; 537 T; 0 other;

Query Match 41.3%; Score 332; DB 21; Length 2027;
 Best Local Similarity 78.3%; Pred. No. 4.1e-96;
 Matches 398; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 1 atggccttggagagcctcttgagcaagttggagccttggagagattcagatgcttcac 60
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 270 atggccttggagagcctcttgagcaagttggagccttggagagattcagatgcttcac 329
 QY 61 ctggtttttttcttcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 120
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 330 actgttttctcctcaatctctgtgtgacacactcctccttattatgtcggagaaacttc 389
 QY 121 gctcagcattccctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 180
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 390 actgcatctactcctgcatcgtcgtggttccacatcctcctgagacatgctcctcct 449
 QY 181 ggtatgaactggaatcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 240
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 450 gaaatgacactggggcctcagcagcagatgacacttggaaatcctcctcctcctcctcctcctcctcct 509
 QY 241 tcaaatctaggcagagagatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 300
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||


```
QY 181 ggtatgaactggaactcctcagtgagaatgcccctcttgagaatctctatccactagac 240
Db 228 ----cgtttccacaacaacatgaccccccaaggccctctctgacacatctccatcccgcaagc 283
QY 241 tcaaatctgagggccagagaagtgctcgctctgttccatccccagtgagcagctctctac 300
Db 284 cccaaccagggggcccaacacagtgccgcgcttccgcagccacagtgagcagctcttgac 343
QY 301 ctgaatgggaactatccacaagcaagtgaggcagacacagaaccctgtgtgatagctgg 360
Db 344 cccaatgccacggccacacagctggaagcgaagctgacacgagccgtgtgtgacgycgtg 403
QY 361 gtatatgtatcaagctacttcccttcgaccatgtgactaagtgggaacctggtatgtgat 420
Db 404 gtctatgacccgacgctcttcaacctccaccatcgtggccaagtggaacctgtgtgacgc 463
QY 421 tatcagtcactgaaatcagtggttcaatccctactctgaactggaatgctgtggaagc 480
Db 464 tcccagggcttgaagcccccctaagcccaatccttcatgttcgsggataccctgtgsgctcc 523
QY 481 atcataagtgcccatgtctcacaagaagt 508
Db 524 ttatctggggcctcctcctcctaccggt 551
```

Search completed: July 1, 2002, 15:45:51
Job time: 6260 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2002, 12:08:21 ; Search time 2405.18 Seconds

(without alignments)
4511.742 Million cell updates/sec

Title: US-09-674-235-10

Perfect score: 804

Sequence: 1 atggccttgagagactctt.....acacaaagtgcacaaattt 804

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST.*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlin:*
5: em_estlin:*
6: em_estlin:*
7: em_estlin:*
8: em_estlin:*
9: em_estlin:*
10: em_estlin:*
11: em_estlin:*
12: em_estlin:*
13: em_estlin:*
14: em_estlin:*
15: em_estlin:*
16: em_estlin:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	494	61.4	615	AV652088	AV652088 AV652088
2	386	48.0	688	AG103175	AG103175 Pan trogl
3	351	43.7	399	AA680184	AA680184 z11c10.s
4	319.2	39.7	581	B1338656	B1338656 362506 MA
5	307.2	38.2	367	AA682573	AA682573 2120406.S
6	293.6	36.4	820	BG972774	BG972774 602857672
7	292.6	36.4	1054	BF783752	BF783752 602109666
8	289.6	36.0	628	BB661512	BB661512 BB661512
9	289.2	35.4	739	B1102429	B1102429 602889616
10	284.8	35.4	1914	AW107114	AW107114 um18p07.y
11	283.2	35.0	615	BB660759	BB660759 Mus muscu
12	281.2	34.9	974	AA1530049	AA1530049 u188f01.y
13	280.4	34.6	694	BB654082	BB654082 BB654082
14	278.2	34.2	578	AA605334	AA605334 AV605334
15	272.8	33.9	653	BB602405	BB602405 BB602405
16	272.6	33.9	754	BG566618	BG566618 602585464

18	266.2	33.1	628	9	BB660738	BB660738
19	265.8	33.1	807	10	B1332219	B1332219
20	264.8	32.9	492	9	AI663462	AI663462
21	261.4	32.5	608	9	AI746617	AI746617
22	261.2	32.5	741	12	AQ237446	AQ237446
23	260.4	32.4	658	9	AI316141	AI316141
24	255.6	31.8	637	9	BB661317	BB661317
25	253.8	31.6	714	9	AA033971	AA033971
26	253.2	31.5	505	10	B1143310	B1143310
27	253	31.5	463	9	AA259000	AA259000
28	250.6	31.2	575	9	AA475229	AA475229
29	246.2	30.6	630	9	BB661292	BB661292
30	243.4	30.3	445	9	AA475431	AA475431
31	242.6	30.2	490	9	AA106794	AA106794
32	240.4	29.9	612	9	BB637428	BB637428
33	234.2	29.1	902	10	BF785293	BF785293
34	233	29.0	811	10	B1328477	B1328477
35	230.8	28.7	451	9	AA012307	AA012307
36	230.8	28.7	612	9	AA611026	AA611026
37	230.4	28.7	915	10	BF788430	BF788430
38	229.2	28.5	607	9	AA195697	AA195697
39	229.2	28.4	570	9	AA107600	AA107600
40	226	28.1	576	9	AI315917	AI315917
41	214.2	26.6	730	12	BH257234	BH257234
42	207.6	25.8	617	9	BB660883	BB660883
43	200.8	25.0	571	9	BE235793	BE235793
44	198.8	24.7	512	9	AI882126	AI882126
45	197.4	24.6	901	10	B1763962	B1763962

ALIGNMENTS

RESULT 1
AV652088 615 bp mRNA linear EST 15-JAN-2002
LOCUS AV652088 GLC Homo sapiens CDNA clone GLOCWF10 3', mRNA sequence.
DEFINITION AV652088
ACCESSION AV652088
VERSION AV652088.1 GI:9873102
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Xu,X., Huang,J., Xu,Z., Qian,B., Zhu,Z., Yan,Q., Cai,T., Zhang,X.,
I (bases 1 to 615)
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

TITLE
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)

JOURNAL
MEDLINE
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzgchgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
location/Qualifiers
1..615
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GLOCWF10"
/clone_lib="GLC"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)

Wagi, K., Fujisake, S., Inoue, K., Togawa, Y., Iwawa, M., Ohara, E., Matabiki, M., Yoneda, Y., Ishikawa, T., Osawa, K., Tanaka, T., Matsura S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

KIKUMI integrated sequence analysis (RISA) system-364-Format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)

Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y., and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)

Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanahe, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. *Mamm. Genome*, 12, 673-677 (2001). Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES	Location/Qualifiers
source	1. . 628

BASE COUNT	156 a	160 c	139 g	171 t	2 others
ORIGIN					

Query Match	36.0%;	Score 289.6;	DB 9;	Length 628;
Best Local Similarity	72.0%;	Pred. No. 5e-64;		
Matches 376;	Conservative 0;	Mismatches 146;	Indels 0;	Gaps 0

QY 1 atgagcccttggagagcctttgagtcgaatgttagagcccttggagagatttcagatgcttcac 60
Db 85 ATGGCCTTTGAGGACCTTATATTAATTCAAATTGGACAGCTTGGGAGAGTTCACAAATCTTCAC 144
QY 61 ctggttttattcttccctccatacgttatataaccatactactgtagaagacttt 120
Db 145 ATGATATTTGGTTAATCTGCAATGCCATGCCCTCTGGGCTCCTCAACCTCTTTGGAGAACTTC 204
QY 121 gctcagagcattccctggtcgcacgttgcctgggttcacacatgctcgagcaataatactgactc 180
Db 205 ACTGAGGCATTCTCTAGTCATCGCTGCTGGGTTCCCCACATCTTGCAATGACACAGCCCTCT 264
QY 181 ggtatgaactgtaactcctcagtgaagatgcccctcttggaaactctatccaccagac 240
Db 265 GATATATGGCACTAGATCCTGAGCCAGAGATGACCTCCTGAGAGATCTCCATCCGCCCTTGAC 324
QY 241 tcaatctgagagccagagaagtgtcgtcgctttgttcataccccaagtggcagctcttcac 300

Db	325	TTCAACCTGGAGACCGGATTAATGGCGTGCTATATTCAACACAGTGGCATCTCTTCAT	388
Oy	301	CTGAATGGGACATCTCCAGCAACAAGTGAAGCAGACACAGAAACCCCTGTGTGAATGCTG	366
Db	385	CTGAATGGCACCCTTCCCACTGTAAACAAGCAACAACAGAGCCCTGTGGAGAGCCTCG	444
Oy	361	GTATATGCAAGTACTTCCTCCCTGAGCAATGTGAATGATGGAGACCTGTATGAT	420
Db	445	GTGATGACCAACACCCCTCTCTCCACACACTGACCCAGTGGAGCTGGTGCTGGGA	504
Oy	421	TATCAATCAAGAACTCAATGCTTACTTCACTGATGATGATGATGATGATGATGATGAT	480
Db	505	TCTCAGGACAGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	564
Oy	481	ATCATAGTGGCATGCTCAAGAAAGTGTGCTGTGTAATCT	522
Db	565	ATCATAGTGGCATGCTCAAGAAAGTGTGCTGTGTAATCT	606

RESULT	9
B1102429	
LOCUS	866 bp mRNA linear EST 26-JUN-2001
DEFINITION	G0288961GFI NCL CGAP_Kid14 Mus musculus cDNA clone IMAGE:5044850
ACCESSION	B1102429
VERSION	B1102429.1 GI:14553322
KEYWORDS	EST.
SOURCE	house mouse.

REFERENCE 1 (bases 1 to 866)
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC).
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC).
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.

FEATURES	Location/Qualifiers
source	1. .866

BASE COUNT	202 a	206 c	189 g	268 t	1 others
ORIGIN					

Query Match	36.0%;	Score 289.2;	DB 10;	Length 866;
Best Local Similarity	72.9%;	Pred. No. 6.5e-64;		
Matches 372;	Conservative	0;	Mismatches 138;	Indels 0;
			Gaps	0

QY 1 atggcctttgagagcgcctcttgatccaagttggagagccttggagatttcagatcttcac 60
 |||||
 Db 173 ATGGCCTTTTCAGGAACCTCCTGAACAAGTTGGAAGCCTAGGAAGATTCAGATCCTTCAG 232
 |||||
 QY 61 ctgagtttatctctccctctccatgttatatccctcatataacgtcaggagaactt 120
 | ||||| ||||| | ||||| |||||

TITLE
JOURNAL

Direct Submission
Submitted (31-Oct-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILN)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: villalobcm.tmc.edu
Villalon, D.K., Luna, R.A., Hale, S.M., Huiy, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny, D.M., Gibbs, R.A.

REMARK
COMMENT

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LILN at: <http://image.liln.gov>
Series: IRAC Plate: 31 Row: e Column: 7
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein
This clone has the following problem: frame shifted.

FEATURES
source

1..1914
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:4236791"
/issue_type="kidney, normal, 5 month old male mouse."
/clone_lib="MCLCGAP_Kid14"
/lab_host="DH10b"
/note="Vector: PCMV-SPORT6"
BASE COUNT 560 a 376 c 385 g 593 t
ORIGIN

Query Match 35.2% Score 283.2; DB 11; Length 1914;
Best Local Similarity 73.6%; Pred. No. 2.5e-62;
Matches 374; Conservative 0; Mismatches 133; Indels 1; Gaps 1;

QY 1 atggccttgagagccttgagcaagttggagccttgaggagattcagatgctcat 60
DB 156 ATGGCCTTGAGAGCCTTGAGCAAGTTGGAGAGCTTGAAGATTCCAGATCTTCAG 215
QY 61 ctggttttatctccctctcctcatgttattcaatcccatatactctagagaactt 120
DB 216 ATCTTTTGTCTCTCTCTCTAGGCCCTTGTGTACTCATATTACATGAGAGACTTT 275
QY 121 gctgagcattcctgctgctgctgagtcagagccttgaggaattcagatgctcat 180
DB 276 ACTGAGCATTCCCAATCATGCTGCTGCGGCCCATCTTGCACATGACACATGCTCT 335
QY 181 ggtatgaactgaactcctcgaagtaagatgaccttgagagatctatcccatagac 240
DB 336 GATATATGACATGAGATCTCTGAGCAAGATGACCTCTGAGAGATCTCCATCCCTTGAT 395
QY 241 tcaatctgagagcagagaagtgtcgtcttgcattccatcccatgagcagcttcaac 300
DB 396 TCCAACTGAGACTGAGATGAATGTCGTG-TTTGCCCAACCAAGTGCATCTCTTCAT 454
QY 301 ctgaatgagacatcacacacacagtgaggcagacacacacacacacacacacacac 360
DB 455 TTGAATGGCACTTCTCCAAATGAGACAGACAGACACTGAGCCCTGTGTGATGGTGG 514
QY 361 gtaataataaagactcctcctcgaacattgtgactaagtgagcctggtatgtat 420
DB 515 GTGTATGACAGAGACACTTCTTCTTACCATGTGACTGATGAGACCTGCTGTGTGA 574
QY 421 tatcagcactggaatcagtggttcaatctcctactctgactgagatgctggtggagac 480

DB 575 TCTCAGGCACTGATTTCTGTGCTAAATTTTCATTCATGATGCGCTATTATAGGGCT 634
QY 481 atcatagtgccatgctcagacagagt 508
DB 635 ATCATATGTGGCATTTGTCAGACAGCT 662

RESULT 12
LOCUS BB660759
DEFINITION BB660759 RIKEN full-length enriched, 0 day neonate kidney mus
ACCESSION musculus cDNA clone D63003H14 5', mRNA sequence.
VERSION BB660759.1 GI:16494539
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 615)
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,
Hiramoto, K., Horii, F., Ishii, Y., Ito, M., Kawai, J., Kono, H., Kouda,
M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,
Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, R., Sano, H., Sasaki,
Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toyota, T.,
Muramatsu, M., and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished (2001)
Contact: Yoshinori Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/,
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,
M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1611-1630 (2000)
wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura,
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and
Hayashizaki, Y.

TITLE
JOURNAL
COMMENT

RIKEN Integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multichannel sequencer. Genome Res. 10
(11), 1757-1771 (2000)
Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara,
Y., and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Alizawa,
K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., and
Hayashizaki, Y.
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for
further details.
e mouse tissues.

FEATURES
source

Location/Qualifiers
1..615
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="D63003H14"
/clone_lib="RIKEN full-length enriched, 0 day neonate
kidney"
/tissue_type="kidney"
/dev_stage="0 day neonate"

[illegible]

Db 683 ATCATATGTGGC 694

RESULT 15

AV605334

LOCUS AV605334 Bos taurus kidney fetus Bos taurus cDNA clone EIKI028B10

DEFINITION 578 bp mRNA linear EST 28-NOV-2001

5', mRNA sequence.

AV605334

ACCESSION

AV605334

VERSION

AV605334.1

KEYWORDS

EST.

SOURCE

cow.

ORGANISM

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

REFERENCE

1 (bases 1 to 578)

AUTHORS

Takasuga,A., Hirotsune,S., Itoh,R., Jitchono,A., Suzuki,H., Aso,H.

TITLE

Establishment of a high throughput EST sequencing system using poly(A) tail-removed cDNA libraries and determination of 36,000 bovine ESTs

JOURNAL

Nucleic Acids Res. 29 (22), E108 (2001)

MEDLINE

21570554

COMMENT

Contact: Yoshikazu Sugimoto

FEATURES

Animal Genetics Division

LOCATION

Shirakawa Institute of Animal Genetics

ODAKURA

Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan

TEL

Tel: 81-248-25-5641

FAX

Fax: 81-248-25-5725

EMAIL

Email: kazusugie@ocn.ne.jp

SINGLE PASS SEQUENCING

This clone was obtained from a polyA-deleted cDNA library.

LOCATION/QUALIFIERS

1..578

ORGANISM

/organism="Bos taurus"

DB XREF

/db_xref="taxon:9913"

CLONE

/clone="EIKI028B10"

ISSUE

/issue="Bos taurus kidney fetus"

DEV STAGE

/dev_stage="fetus"

LAB HOST

/lab_host="DH10B"

NOTES

/note="Vector: pZL1; Site1: SalI; Site2: NotI; Poly A was deleted from a NotI site"

BASE COUNT

117 a 164 c 137 g 160 t

ORIGIN

Query Match

BEST LOCAL SIMILARITY

34.2%; Score 274.8; DB 9; Length 578;

MATCHES

367; Conservative 0; Mismatches 107; Indels 18; Gaps 1;

DB

17 tcttgagtcagggtgagccttgagagatcagatgcttcacgtgttattatcttc 76

DB

12 TCCTTAATGAAGTTGGGCTGGGAAATTCACAGATCCTTCAGATGTTTAACTCTTC 71

DB

77 cctctccatgattatcatcccatatatactgctagagaacttgcgcagccattctg 136

DB

72 CTCTGTGTGGATGACGACCTGTCATATATCTGTGGAGAAATTTCACCTGCTGCCATTCTCG 131

DB

137 gtcatcgctggtggtcacaatgctgagacaataatactgagatcgtgtaataagaaatgaa 196

DB

132 GTGATGCTGCTGTGGTCTACTCTCTGA-----TAATGCCACTGAGAG 173

DB

197 tctcagtgagagatgacctcttgagaatctctatcccaactagaccataatctgagagccag 256

DB

174 TCCACAGCCCTGATGCTCTCTGAGAAATCTCAGATCCACCTGATTCAAACTTAAAGCCAG 233

DB

257 agaagtgctgctgttgctcatcccccagtgagagctctcttaactgaatgagactatcc 316

DB

234 AGAAGTGTCATGCTCTCTCCACCCGAGTGCGACAGCTCTTCACCTGGAATAGGACCTTCG 293

QY 317 acagcacagtgagagcagacacagaccctgtgtgagtgctgggtatgatgataaagct 376
 Db 294 CCAACATGACACTGGCGCTGGACACGAGCCCTGTATGAGAGCGGTGATGACACAGAGCT 353
 QY 377 actcccttcgaccatttgactaagtgagaccctgtgtgattatgactgaactgaat 436
 Db 354 TGTCTCTCTCCACCATCTGCTGACTGAGTGGACCTCATATGTGATACACGATACAGAAAT 413
 QY 437 cagtggtcattctctactcttgactggaatgctgtgtgagagcactacatagtgagcag 496
 Db 414 CAGTGTTCAAGTTGTATTCATGAGGCTGGAATGACAGTGGAGGCTTCGTATACGGGCATC 473
 QY 497 tctcagacaggt 508
 Db 474 TCTCAGACAGGT 485

Search completed: July 1, 2002, 14:44:48
 Job time: 9387 sec

us-09-674-235-10 rmj

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/647,397
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Gates, Edward R.
REGISTRATION NUMBER: 31,616
REFERENCE/DOCKET NUMBER: B0801/7048
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2102 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: Internal
ORIGINAL SOURCE:
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: 120..1733
US-08-647-397-1

```

```

1 ADDRESS: Foley & Lardner
2 STREET: 1800 Diagonal Road, Suite 500
3 CITY: Alexandria
4 STATE: VA
5 COUNTRY: USA
6 ZIP: 22313-0299
7 COMPUTER READABLE FORM:
8 MEDIUM TYPE: Floppy disk
9 COMPUTER: IBM PC compatible
10 OPERATING SYSTEM: PC-DOS/MS-DOS
11 SOFTWARE: Patent In Release #1.0, Version #1.25
12 CURRENT APPLICATION DATA:
13 APPLICATION NUMBER: US/08/232,463
14 FILING DATE:
15 CLASSIFICATION: 435
16 PRIOR APPLICATION DATA:
17 APPLICATION NUMBER: US/07/935,313
18 FILING DATE:
19 APPLICATION NUMBER: EP 91 114 300.6
20 FILING DATE: 26-AUG-1991
21 ATTORNEY/AGENT INFORMATION:
22 NAME: BENT, Stephen A.
23 REGISTRATION NUMBER: 29,768
24 REFERENCE/DOCKET NUMBER: 30472/114 IMMU
25 TELECOMMUNICATION INFORMATION:
26 TELEPHONE: (703)836-9300
27 TELEFAX: (703)683-4109
28 TELE: 899149
29 INFORMATION FOR SEQ ID NO: 14:
30 SEQUENCE CHARACTERISTICS:
31 LENGTH: 7218 base pairs
32 TYPE: nucleic acid
33 STRANDEDNESS: single
34 TOPOLOGY: linear
35 IMMEDIATE SOURCE:
36 CLONE: PTgpc-Fls
37 US-08-232-463-14

```

Query Match	6.6%	Score 53, 2:	DB 2:	Length 2102;
Best Local Similarity	56.5%	Pred. No. 2.5e-07;		
Matches 122: Conservative	0			

Query Match 5.8%; Score 46.6; DB 1; Length 7218;
Best Local Similarity 5.7%; Pred. No. 6.8e-05;
Matches 22; Conservative 203; Mismatches 162; Indels 0; Gaps 0.

	Conservative	V	Mismatches	88	Indels	6	Gaps	1
QY	293	ttcttcacctaagatgagactatccacagacaaagtgaggagcagacagaaacctgtgtg	352					
Db	349	TGCATCTGCCAAAGCGCAGTCTTCCCATGTGACACCCAGGGGCCACCGAGCATCTTGG	408					
QY	353	atggctgtggtatgatcatcaaaagctactcccttcgaccattgtgactaagtggagacctg	412					
Db	409	ATGGCTGGATCT-----ACAACAGCACACAGAACACCATTTGTGACAGAGTGGGACTTGG	462					
QY	413	tatgtgttatcagtaactgaatcaatcagtggttcaattcctaactctgactggaatcctg	472					
Db	463	TATCGGGCTCCAAACAACTGAGAGAGATGGCAGCAGTCACTTATGCGCAGGTATACGG	522					
QY	473	tgggaggcatacatggtggtgcacatgtctccagaagaagt	508					
Db	523	TTGGAGGACCTGTCTTTTGGAGAACTGTCAACACAGGT	558					

[illegible]

RESULT 3
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOMLPX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:

```
QY      775 aatgattgctacacaaagtaccaa 801
::: : : : : : : : : :
Db      1124 RRRRRRRRRRRRRRRRRRRRRRR 1098
```

```

US-08-232-463-14
: Sequence 14, Application US/08232463
: Patent No. 5670367
:
: GENERAL INFORMATION:
: APPLICANT: DORNER, F.
: APPLICANT: SCHEIFLINGER, F.
: APPLICANT: FALKNER, F. G.
: TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 1800 Diagonal Road, Suite 500
: CITY: Alexandria
: STATE: VA
: COUNTRY: USA
: ZIP: 22313-0299
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/232,463
:
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/935,313
: FILING DATE:
: APPLICATION NUMBER: EP 91 114 300.6
: FILING DATE: 26-AUG-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 30472/114 IMMU
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703)836-9300
: TELEFAX: (703)683-4109
: TELEX: 899149
:
: INFORMATION FOR SEQ ID NO: 14:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 7218 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: CLONE: pTZ9pt-F15
:
US-08-232-463-14

```

[illegible]

```

1  RESULT 5
2  US-07-906-871-15//C
3  ; Sequence 15, Application US/07906871
4  ; Patent No. 5340739
5  ;
6  ; GENERAL INFORMATION:
7  ; APPLICANT: Stevens, Richard L.
8  ; APPLICANT: Avraham, Shalom
9  ; TITLE OF INVENTION: HEMATOPOIETIC CELL SPECIFIC
10 ; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATORY ELEMENTS OF SERGLYCIN AND USES
11 ; TITLE OF INVENTION: THEREOF
12 ; NUMBER OF SEQUENCES: 18
13 ;
14 ; CORRESPONDENCE ADDRESS:
15 ; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
16 ; STREET: 1225 Connecticut Avenue, N.W., Suite 300
17 ; CITY: Washington
18 ; STATE: DC
19 ; COUNTRY: USA
20 ; ZIP: 20036
21 ;
22 ; COMPUTER READABLE FORM:
23 ; MEDIUM TYPE: Floppy disk
24 ; COMPUTER: IBM PC compatible
25 ; OPERATING SYSTEM: PC-DOS/MS-DOS
26 ; SOFTWARE: Patentln Release #1.0, Version #1.25
27 ;
28 ; CURRENT APPLICATION DATA:
29 ; APPLICATION NUMBER: US/07/906, 871
30 ; FILING DATE: 19920103
31 ;
32 ; CLASSIFICATION: 435
33 ;
34 ; PRIOR APPLICATION DATA:
35 ; APPLICATION NUMBER: US/07/816, 289
36 ; FILING DATE: 03 JAN 1992
37 ;
38 ; PRIOR APPLICATION DATA:
39 ; APPLICATION NUMBER: US/07/535, 544
40 ; FILING DATE: 18-JAN-1991
41 ;
42 ; PRIOR APPLICATION DATA:
43 ; APPLICATION NUMBER: PCT/US89/03051
44 ; FILING DATE: 13-JUL-1989
45 ;
46 ; PRIOR APPLICATION DATA:
47 ; APPLICATION NUMBER: US/07/224, 035
48 ; FILING DATE: 13-JUL-1988
49 ;
50 ; ATTORNEY/AGENT INFORMATION:
51 ; NAME: Cimbalia, Michele A
52 ; REGISTRATION NUMBER: 33, 851
53 ; REFERENCE/DOCKET NUMBER: 0627, 2830004
54 ;
55 ; TELECOMMUNICATION INFORMATION:
56 ; TELEPHONE: (202)833-7533
57 ; TELEFAX: (202)833-8716
58 ;
59 ; INFORMATION FOR SEQ ID NO: 15:
60 ; SEQUENCE CHARACTERISTICS:
61 ; LENGTH: 17327 base pairs
62 ; TYPE: NUCLEIC ACID
63 ; STRANDEDNESS: both
64 ; TOPOLOGY: linear
65 ;
66 ; MOLECULE TYPE: DNA
67 ;
68 ; FEATURE:
69 ; NAME/KEY: exon
70 ; LOCATION: 621..753
71 ;
72 ; FEATURE:
73 ; NAME/KEY: intron
74 ; LOCATION: 754..9596
75 ;
76 ; FEATURE:
77 ; NAME/KEY: exon
78 ; LOCATION: 9597..9744
79 ;
80 ; FEATURE:
81 ; NAME/KEY: intron
82 ; LOCATION: 9745..16396
83 ;
84 ; FEATURE:
85 ; NAME/KEY: exon
86 ; LOCATION: 16397..17327
87 ;
88 ; US-07-906-871-15

```

Query Match 4.38; Score 34.2; DB 1; Length 17327;

Best Local Similarity 51.7%; Pred. No. 1.2;
Matches 78; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 527 ggttggtgtaatacaccataactagatgaggtcctaagcacttaagaagtgcac 586
DB 14758 ggttcgcgagagaaacacgctgtaagagagggccgacagccgctcacac 14699
QY 587 gacaaatgaaataaagaatgctgaagaacccctgaacataagatggttaagatccacaa 646
DB 14698 CAACATGTTAAATACCTGACAGCAAGGAAAAAGCACAACAGAGATGAGAGAAACAGGA 14639
QY 647 tgcagagagagctgagtgatgagcagaccacaa 677
DB 14638 TAAGCAGAGGCGATGAGAGCTCAAGAAAAA 14608

RESULT 6
US-09-081-320-1
; Sequence 1, Application US/09081320
; Patent No. 6093544
; GENERAL INFORMATION:

APPLICANT: Gonsalves, Dennis
APPLICANT: Meng, Baozhong
TITLE OF INVENTION: RUPESTRIS STEM PITTING ASSOCIATED VIRUS
NUMBER OF INVENTION: NUCLEIC ACIDS, PROTEINS, AND THEIR USES
CORRESPONDENCE ADDRESS: 54
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/081,320
FILING DATE:
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/047,147
FILING DATE: 20-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/069,902
FILING DATE: 17-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/17222
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600

INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8743 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-081-320-1

Query Match 4.1%; Score 33.2; DB 3; Length 8743;
Best Local Similarity 50.0%; Pred. No. 1.7;
Matches 83; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 559 ggcttaagcacttaagaatgacgcacaaatgataagaatgctgaagaacc 618
DB 7764 GGCTGAATGGCAATGTAATGGAAACTCCCGGTGAATCAATGAGGCTTTGAGACC 7823

QY 619 ctgaacatagaggtgtaagatccacatgacagagagctgtagcacagaccacaa 678
DB 7824 CGGCTAAATCGCTGGAGTTAGCTAGAGCTCAAAAGACGCCGAGAGTTTAAATGACCA 7883
QY 679 actacgtgtgactgtgttcgcgaacccagatagcgtataaaga 724
DB 7884 CTTACTCTCAGTGGCATTTCTTGCCAAACGCAAGAGATTTATAGAGA 7929

RESULT 7
US-08-658-665-178
; Sequence 178, Application US/08658665
; Patent No. 5997878
; GENERAL INFORMATION:

APPLICANT: Paoletti, Enzo
APPLICANT: Pincus, Steven E.
APPLICANT: Cox, William I.
APPLICANT: Kauffman, Elizabeth K.
TITLE OF INVENTION: Recombinant Foxvirus - Cytomegalovirus,
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis, Morris & Safford, P.C.
STREET: 530 Fifth Avenue
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/658,665
FILING DATE: 05-JUN-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Frommer Esq., William S.
REGISTRATION NUMBER: 25,506
REFERENCE/DOCKET NUMBER: 454310-2720.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)840-3333
TELEFAX: (212)840-0712

INFORMATION FOR SEQ ID NO: 178:
SEQUENCE CHARACTERISTICS:
LENGTH: 5798 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-658-665-178

Query Match 4.0%; Score 32.4; DB 2; Length 5798;
Best Local Similarity 48.0%; Pred. No. 2.5;
Matches 122; Conservative 0; Mismatches 131; Indels 1; Gaps 1;

QY 548 aactagatgaggtcctaagaagcacttagaagaagtgcagcacaatgataaagaatg 607
DB 1904 AATTGGTAGACGAAATAATTTATCTAAACGACATCGTCTAAACGAAATGAGATTA 1963
QY 608 ctgaagaacccctgaacatagaggtgtgaagatccacatgacagagagctgtagcag 667
DB 1964 AAGAGAGATTAGTAACGAAATTAATGGAATAATGTTTCGAAGATTAATGACTA 2023

QY 668 ccacagacaaactctggtgtgactgttccgcgaacccagatgagtaagaatct 727
DB 2024 TACCTACCGATTCCTATGAGATA-TTTTTTAAACCGGTACTAGAGAAAAAGTATCT 2082
QY 728 gtatcctgtatcttctgaagaaaaaatctcaaggaagaaagcataaagtgtgtctaa 787
DB 2083 AAAGCTGTAGATTTTCCAGATGATATTAAAGGAGATGATTAATGCAAAATGGAGATA 2142

QY	788	caaaagtaccaa	801
Db	2143	AAACACGGAGAAA	2156

```

1 RESULT 8
2 US-08-796-101-39
3 ; Sequence 39, Application US/08796101
4 ; Patent No. 6183752
5 ;
6 ; GENERAL INFORMATION:
7 ;
8 ; APPLICANT: EPSTEIN, STEPHEN E.
9 ; APPLICANT: FINKEL, TOREN
10 ; APPLICANT: SPEIR, EDITH
11 ; APPLICANT: ZHOU, YI FU
12 ; APPLICANT: ZHU, JIANNUI
13 ; APPLICANT: ERDIE, LORNE
14 ; APPLICANT: PINCUS, STEVEN
15 ; TITLE OF INVENTION: RESTENOSIS/ATHEROSCLEROSIS DIAGNOSIS,
16 ; TITLE OF INVENTION: PROPHYLAXIS AND THERAPY
17 ;
18 ; NUMBER OF SEQUENCES: 184
19 ;
20 ; CORRESPONDENCE ADDRESS:
21 ; ADDRESSEE: CURTIS, MORRIS & SAFFORD, P.C.
22 ; STREET: 530 FIFTH AVENUE
23 ; CITY: NEW YORK
24 ; STATE: NY
25 ;
26 ; COUNTRY: USA
27 ;
28 ; ZIP: 10036
29 ;
30 ; COMPUTER READABLE FORM:
31 ;
32 ; MEDIUM TYPE: Floppy disk
33 ; COMPUTER: IBM PC compatible
34 ; OPERATING SYSTEM: PC-DOS/MS-DOS
35 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
36 ;
37 ; CURRENT APPLICATION DATA:
38 ; APPLICATION NUMBER: US/08/796,101
39 ; FILING DATE: 05-FEB-1997
40 ;
41 ; CLASSIFICATION: 424
42 ;
43 ; ATTORNEY/AGENT INFORMATION:
44 ;
45 ; NAME: KOWALSKI, THOMAS J.
46 ; REGISTRATION NUMBER: 32,147
47 ; TELECOMMUNICATION INFORMATION:
48 ; TELEPHONE: (212) 840-3333
49 ; TELEFAX: (212) 764-5574
50 ;
51 ; INFORMATION FOR SEQ ID NO: 39:
52 ;
53 ; SEQUENCE CHARACTERISTICS:
54 ;
55 ; LENGTH: 5798 base pairs
56 ;
57 ; TYPE: nucleic acid
58 ;
59 ; STRANDEDNESS: single
60 ;
61 ; TOPOLOGY: linear
62 ;
63 ; MOLECULE TYPE: DNA (genomic)
64 ;
65 ;
66 ;
67 ;
68 ;
69 ;
70 ;
71 ;
72 ;
73 ;
74 ;
75 ;
76 ;
77 ;
78 ;
79 ;
80 ;
81 ;
82 ;
83 ;
84 ;
85 ;
86 ;
87 ;
88 ;
89 ;
90 ;
91 ;
92 ;
93 ;
94 ;
95 ;
96 ;
97 ;
98 ;
99 ;
100 ;
101 ;
102 ;
103 ;
104 ;
105 ;
106 ;
107 ;
108 ;
109 ;
110 ;
111 ;
112 ;
113 ;
114 ;
115 ;
116 ;
117 ;
118 ;
119 ;
120 ;
121 ;
122 ;
123 ;
124 ;
125 ;
126 ;
127 ;
128 ;
129 ;
130 ;
131 ;
132 ;
133 ;
134 ;
135 ;
136 ;
137 ;
138 ;
139 ;
140 ;
141 ;
142 ;
143 ;
144 ;
145 ;
146 ;
147 ;
148 ;
149 ;
150 ;
151 ;
152 ;
153 ;
154 ;
155 ;
156 ;
157 ;
158 ;
159 ;
160 ;
161 ;
162 ;
163 ;
164 ;
165 ;
166 ;
167 ;
168 ;
169 ;
170 ;
171 ;
172 ;
173 ;
174 ;
175 ;
176 ;
177 ;
178 ;
179 ;
180 ;
181 ;
182 ;
183 ;
184 ;
185 ;
186 ;
187 ;
188 ;
189 ;
190 ;
191 ;
192 ;
193 ;
194 ;
195 ;
196 ;
197 ;
198 ;
199 ;
200 ;
201 ;
202 ;
203 ;
204 ;
205 ;
206 ;
207 ;
208 ;
209 ;
210 ;
211 ;
212 ;
213 ;
214 ;
215 ;
216 ;
217 ;
218 ;
219 ;
220 ;
221 ;
222 ;
223 ;
224 ;
225 ;
226 ;
227 ;
228 ;
229 ;
230 ;
231 ;
232 ;
233 ;
234 ;
235 ;
236 ;
237 ;
238 ;
239 ;
240 ;
241 ;
242 ;
243 ;
244 ;
245 ;
246 ;
247 ;
248 ;
249 ;
250 ;
251 ;
252 ;
253 ;
254 ;
255 ;
256 ;
257 ;
258 ;
259 ;
260 ;
261 ;
262 ;
263 ;
264 ;
265 ;
266 ;
267 ;
268 ;
269 ;
270 ;
271 ;
272 ;
273 ;
274 ;
275 ;
276 ;
277 ;
278 ;
279 ;
280 ;
281 ;
282 ;
283 ;
284 ;
285 ;
286 ;
287 ;
288 ;
289 ;
290 ;
291 ;
292 ;
293 ;
294 ;
295 ;
296 ;
297 ;
298 ;
299 ;
300 ;
301 ;
302 ;
303 ;
304 ;
305 ;
306 ;
307 ;
308 ;
309 ;
310 ;
311 ;
312 ;
313 ;
314 ;
315 ;
316 ;
317 ;
318 ;
319 ;
320 ;
321 ;
322 ;
323 ;
324 ;
325 ;
326 ;
327 ;
328 ;
329 ;
330 ;
331 ;
332 ;
333 ;
334 ;
335 ;
336 ;
337 ;
338 ;
339 ;
340 ;
341 ;
342 ;
343 ;
344 ;
345 ;
346 ;
347 ;
348 ;
349 ;
350 ;
351 ;
352 ;
353 ;
354 ;
355 ;
356 ;
357 ;
358 ;
359 ;
360 ;
361 ;
362 ;
363 ;
364 ;
365 ;
366 ;
367 ;
368 ;
369 ;
370 ;
371 ;
372 ;
373 ;
374 ;
375 ;
376 ;
377 ;
378 ;
379 ;
380 ;
381 ;
382 ;
383 ;
384 ;
385 ;
386 ;
387 ;
388 ;
389 ;
390 ;
391 ;
392 ;
393 ;
394 ;
395 ;
396 ;
397 ;
398 ;
399 ;
400 ;
401 ;
402 ;
403 ;
404 ;
405 ;
406 ;
407 ;
408 ;
409 ;
410 ;
411 ;
412 ;
413 ;
414 ;
415 ;
416 ;
417 ;
418 ;
419 ;
420 ;
421 ;
422 ;
423 ;
424 ;
425 ;
426 ;
427 ;
428 ;
429 ;
430 ;
431 ;
432 ;
433 ;
434 ;
435 ;
436 ;
437 ;
438 ;
439 ;
440 ;
441 ;
442 ;
443 ;
444 ;
445 ;
446 ;
447 ;
448 ;
449 ;
450 ;
451 ;
452 ;
453 ;
454 ;
455 ;
456 ;
457 ;
458 ;
459 ;
460 ;
461 ;
462 ;
463 ;
464 ;
465 ;
466 ;
467 ;
468 ;
469 ;
470 ;
471 ;
472 ;
473 ;
474 ;
475 ;
476 ;
477 ;
478 ;
479 ;
480 ;
481 ;
482 ;
483 ;
484 ;
485 ;
486 ;
487 ;
488 ;
489 ;
490 ;
491 ;
492 ;
493 ;
494 ;
495 ;
496 ;
497 ;
498 ;
499 ;
500 ;
501 ;
502 ;
503 ;
504 ;
505 ;
506 ;
507 ;
508 ;
509 ;
510 ;
511 ;
512 ;
513 ;
514 ;
515 ;
516 ;
517 ;
518 ;
519 ;
520 ;
521 ;
522 ;
523 ;
524 ;
525 ;
526 ;
527 ;
528 ;
529 ;
530 ;
531 ;
532 ;
533 ;
534 ;
535 ;
536 ;
537 ;
538 ;
539 ;
540 ;
541 ;
542 ;
543 ;
544 ;
545 ;
546 ;
547 ;
548 ;
549 ;
550 ;
551 ;
552 ;
553 ;
554 ;
555 ;
556 ;
557 ;
558 ;
559 ;
560 ;
561 ;
562 ;
563 ;
564 ;
565 ;
566 ;
567 ;
568 ;
569 ;
570 ;
571 ;
572 ;
573 ;
574 ;
575 ;
576 ;
577 ;
578 ;
579 ;
580 ;
581 ;
582 ;
583 ;
584 ;
585 ;
586 ;
587 ;
588 ;
589 ;
590 ;
591 ;
592 ;
593 ;
594 ;
595 ;
596 ;
597 ;
598 ;
599 ;
600 ;
601 ;
602 ;
603 ;
604 ;
605 ;
606 ;
607 ;
608 ;
609 ;
610 ;
611 ;
612 ;
613 ;
614 ;
615 ;
616 ;
617 ;
618 ;
619 ;
620 ;
621 ;
622 ;
623 ;
624 ;
625 ;
626 ;
627 ;
628 ;
629 ;
630 ;
631 ;
632 ;
633 ;
634 ;
635 ;
636 ;
637 ;
638 ;
639 ;
640 ;
641 ;
642 ;
643 ;
644 ;
645 ;
646 ;
647 ;
648 ;
649 ;
650 ;
651 ;
652 ;
653 ;
654 ;
655 ;
656 ;
657 ;
658 ;
659 ;
660 ;
661 ;
662 ;
663 ;
664 ;
665 ;
666 ;
667 ;
668 ;
669 ;
670 ;
671 ;
672 ;
673 ;
674 ;
675 ;
676 ;
677 ;
678 ;
679 ;
680 ;
681 ;
682 ;
683 ;
684 ;
685 ;
686 ;
687 ;
688 ;
689 ;
690 ;
691 ;
692 ;
693 ;
694 ;
695 ;
696 ;
697 ;
698 ;
699 ;
700 ;
701 ;
702 ;
703 ;
704 ;
705 ;
706 ;
707 ;
708 ;
709 ;
710 ;
711 ;
712 ;
713 ;
714 ;
715 ;
716 ;
717 ;
718 ;
719 ;
720 ;
721 ;
722 ;
723 ;
724 ;
725 ;
726 ;
727 ;
728 ;
729 ;
730 ;
731 ;
732 ;
733 ;
734 ;
735 ;
736 ;
737 ;
738 ;
739 ;
740 ;
741 ;
742 ;
743 ;
744 ;
745 ;
746 ;
747 ;
748 ;
749 ;
750 ;
751 ;
752 ;
753 ;
754 ;
7
```

Query Match	Best Local Similarity	48.0%;	Score 32.4;	DB 4;	Length 5798;
Matches 122;	Conservative	0;	Mismatches 131;	Indels 1;	Gaps 1;
QY 548	aactagatgagggttaagcgaccttaagaagaatgcgcacaaatgtaataaagatg				
Db 1904	AATTGGTAGACAGAAAATATATTCTTAACCAAGATACGCTTAACCAAGAAATAGAGATTA				
QY 608	ctgaagaaccctgaacataaggttctgaagatccaccatgcaggaggaagcttgatgcag				
Db 1964	AAAGAACTTTTAAAGCAATTAATTGAAAATACCAGTAATGTTCCAGATATATATAGACTA				
QY 668	cacagaccacaacacacagctctctgactctgttcgcgaaccccaatgctgcgttaaaagatct				
Db 2024	TACCTACCAAGTATCCCTATGAGATG-TTTTTTTAAACCGGTACTAAGAGAAATATATCT				
QY 728	gtatctctgtaatttttgaagaaaaaatctcaaggaaagagcataaataatgctctaca				
Db 2083	AAAGGTGATGATTTTTCAGAAATGATATTTAAGGACAGATGATATTAGCAAAATGGCAATA				

QY	788	caaaagtgaccaa	801
Db	2143	AAACACGAGAAAA	2156

```

1 RESULT
2 9
3 US-07-923-095-1
4 ; Sequence 1, Application US/07923095
5 ; Patent No. 5401652
6 ;
7 GENERAL INFORMATION:
8 APPLICANT: Zhai Dr., Patricia T.
9 APPLICANT: Zhai Dr., Mohammad R.
10 TITLE OF INVENTION: Nucleic Acid Sequence
11 TITLE OF INVENTION: Receptor
12 NUMBER OF SEQUENCES: 2
13 CORRESPONDENCE ADDRESS:
14 ADDRESSEE: American Cyanamid Company
15 STREET: 1937 West Main Street
16 CITY: Stamford
17 STATE: CT
18 COUNTRY: U.S.A.
19 ZIP: 06904-0060
20
21 COMPUTER READABLE FORM:
22 MEDIUM TYPE: Floppy disk
23 COMPUTER: IBM PC compatible
24 OPERATING SYSTEM: PC-DOS/MS-DOS
25 SOFTWARE: patentin Release #1.0, Version 1.0
26 CURRENT APPLICATION DATA:
27 APPLICATION NUMBER: US/07/923, 095
28 FILING DATE: 19920730
29 CLASSIFICATION: 435
30 ATTORNEY/AGENT INFORMATION:
31 NAME: Lowmey Dr., Karen A.
32 REGISTRATION NUMBER: 31274
33 REFERENCE/DOCKET NUMBER: 31619-00
34 TELECOMMUNICATION INFORMATION:
35 TELEPHONE: 203-321-2361
36 TELEFAX: 203-321-2371
37 TELEX: 710-474-4059
38 INFORMATION FOR SEQ ID NO: 1:
39 SEQUENCE CHARACTERISTICS
40 LENGTH: 1730 base pairs
41 TYPE: NUCLEIC ACID
42 STRANDEDNESS: single
43 TOPOLOGY: linear
44 MOLECULE TYPE: DNA
45 HYPOTHETICAL: NO
46 ANTI-SENSE: NO
47 ORIGINAL SOURCE:
48 ORGANISM: pig
49 FEATURE:
50 NAME/KEY: CDS
51 LOCATION: 140..1456
52
53 US-07-923-095-1

```

[illegible]

RESULT 10
US-08-229-511-1
; Sequence 1, Application US/08229511
; Patent No. 5591824
; GENERAL INFORMATION:
; APPLICANT: Ziai, Mohammad Reza
; APPLICANT: Sokol, Patricia Tyson
; APPLICANT: Chandra, Manik
; TITLE OF INVENTION: Apamin Binding Protein/Receptor
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: 2 Militia Drive
; CITY: Lexington
; STATE: Ma
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/229,511
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/922,307
; FILING DATE: 30-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/923,095
; FILING DATE: 30-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Alice O.
; REGISTRATION NUMBER: 33,542
; REFERENCE/DOCKET NUMBER: ACC92-08A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1730 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 140..1456
; US-08-229-511-1

Query Match 3.9%; Score 31.6; DB 1; Length 1730;
Best Local Similarity 52.2%; Pred. No. 2.4;
Matches 70; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
QY 343 ccctgtgtgtagtgcgtgggtatataatgaatcactccctcgaccatgtgactaaag 402
DB 1529 CCCTTATGCTGGGCTTGTGTTTAAAGCCAGCAGCAATGGACACACTTACTCTAG 1588
QY 403 tgggacctgtgtatgtatatacagtcacgtgaatcagtggttcaattcctactctgact 462
DB 1589 TGGGAGATGCAAGATGATGATACAGGGGGGTGTTAGGAATATTTCTAAGTTTCCACCT 1648
QY 463 ggaatcgtgtggg 476
DB 1649 TGAATGCTGATGG 1662

RESULT 11
US-08-314-979-1
; Sequence 1, Application US/08314979
; Patent No. 5607843
; GENERAL INFORMATION:

APPLICANT: Sokol Dr., Patricia T.
APPLICANT: Ziai Dr., Mohammad R.
TITLE OF INVENTION: Nucleic Acid Sequence Encoding an Apamin
Patent No. 5607843
TITLE OF INVENTION: Receptor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: 1937 West Main Street
CITY: Stamford
STATE: CT
COUNTRY: U.S.A.
ZIP: 06904-0060
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/314,979
FILING DATE: 29-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/923,095
FILING DATE: 30-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lowrey Dr., Karen A.
REGISTRATION NUMBER: 31274
REFERENCE/DOCKET NUMBER: 31619-00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 203-321-2361
TELEFAX: 203-321-2971
TELEX: 710-474-4059
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1730 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Pig
FEATURE:
NAME/KEY: CDS
LOCATION: 140..1456
US-08-314-979-1

Query Match 3.9%; Score 31.6; DB 1; Length 1730;
Best Local Similarity 52.2%; Pred. No. 2.4;
Matches 70; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
QY 343 ccctgtgtgtagtgcgtgggtatataatgaatcactccctcgaccatgtgactaaag 402
DB 1529 CCCTTATGCTGGGCTTGTGTTTAAAGCCAGCAGCAATGGACACACTTACTCTAG 1588
QY 403 tgggacctgtgtatgtatatacagtcacgtgaatcagtggttcaattcctactctgact 462
DB 1589 TGGGAGATGCAAGATGATGATACAGGGGGGTGTTAGGAATATTTCTAAGTTTCCACCT 1648
QY 463 ggaatcgtgtggg 476
DB 1649 TGAATGCTGATGG 1662

RESULT 12
US-08-436-716-1
; Sequence 1, Application US/08436716
; Patent No. 5652111
; GENERAL INFORMATION:
APPLICANT: Sokol Dr., Patricia T.


```

APPLICANT: Zhai Dr., Mohammad R.
TITLE OF INVENTION: Nucleic Acid Sequence Encoding an Apamin
Patent No. 5652111
TITLE OF INVENTION: Receptor
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: 1937 West Main Street
CITY: Stamford
STATE: CT
COUNTRY: U.S.A.
ZIP: 06904-0060
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/436,716
FILING DATE: 08-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/923,095
FILING DATE: 30-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Lowney Dr., Karen A.
REGISTRATION NUMBER: 31274
REFERENCE/DOCKET NUMBER: 31619-00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 203-321-2361
TELEFAX: 203-321-2971
TELEX: 710-474-4059
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1730 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Pig
FEATURE:
NAME/KEY: CDS
LOCATION: 140..1456
US-08-436-716-1

Query Match 3.9%; Score 31.6; DB 1; Length 1730;
Best Local Similarity 52.2%; Pred. No. 2.4;
Matches 70; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

OY 343 cccgtgtgtagtgcgtggtatgatcaagaagctactcccttgacacttgtagtaag 402
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1529 CCCCTTATGCGTGGGCTTTGTTGTTTAAACCCAGGACAAATGGCACAGCTTACTCAG 1588

OY 403 tgggaaccggtatgtagtattcaatgcactgaataatcaagtgttcaattcctacttgaact 462
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1589 TGGGAGATGCAGATGACTACCAAGGGGCTGTTAGGATATATTCTAAGTTTTCACCT 1648

OY 463 ggaatgcctgttgg 476
||||| ||| |||
DB 1649 TGAATGCTGAGTGG 1662

RESULT 13
PCT-US92-06840-1/c
; Sequence 1, Application PC/TUS9206840
; GENERAL INFORMATION:
; APPLICANT: Shi, Yang
; APPLICANT: Seto, Edward
; APPLICANT: Shenk, Thomas
;

```

```

TITLE OF INVENTION: YY1 TRANSCRIPTION FACTOR AND METHODS OF
NUMBER OF INVENTION: 10
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
STREET: 1180 Avenue of the Americas - 7th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-8403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/06840
FILING DATE: 19920814
CLASSIFICATION:
CLASSIFICATION: AU 1805
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/746,485
FILING DATE: 16-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Dennis Manette
REGISTRATION NUMBER: 30,623
REFERENCE/DOCKET NUMBER: M-12594 CIP (1570-8)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 382-0700
TELEFAX: (212) 382-0888
TELEX: 236925
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2353 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: HeLa cells derived from cervical
TISSUE TYPE: carcinoma
CELL TYPE: tumor cells
CELL LINE: HeLa
IMMEDIATE SOURCE:
LIBRARY: D98/AH-2
CLONE: p14-1 or pY1
FEATURE:
NAME/KEY: CDS
LOCATION: 241..1485
PCT-US92-06840-1

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2002, 14:44:48 ; Search time 2405.18 Seconds
(without alignments)
9567.811 Million cell updates/sec

Title: US-09-674-235-19
Perfect score: 1705
Sequence: 1 aagacgcagcagcctctt.....aatttaaaattatattc 1705

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

1: em_estha:*
2: em_estha:*
3: em_estha:*
4: em_estha:*
5: em_estha:*
6: em_estha:*
7: em_estha:*
8: em_estha:*
9: gb_est1:*
10: gb_est2:*
11: gb_hc:*
12: gb_hc:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	549	32.2	615	9	AV652088
2	545	32.0	779	9	AV652088
3	388	22.8	688	12	AG103175 Pan trogl
4	387	22.7	399	9	AA680184 zllc10.s
5	382	22.6	612	12	AQ280703 CITR1-EL
6	372.8	21.9	474	12	AQ127694 HS-3091-B
7	370	21.7	578	12	AQ280563 CITR1-EL
8	346.2	20.3	367	9	AA682573 z120a06.s
9	336.4	19.7	754	10	BG566618 602585464
10	328.8	19.3	581	10	BI338656 362506 MA
11	314	18.4	741	12	AQ237446 RPK11-69
12	298.2	17.5	820	10	BG972774 602837672
13	293.8	17.2	866	10	BI102429 602889616
14	292.6	17.2	739	9	AM107114 um18h07.Y
15	289.8	17.0	1054	10	BF783752 602109666
16	289.8	17.0	628	9	BB661512 BB661512
17	287.8	16.9	1914	11	BC016496 Mus muscu

18	284	16.7	974	9	AI330049	AI330049 um8f01.Y
19	282.8	16.6	694	9	BB654082	BB654082 BB654082
20	281.4	16.5	615	9	BB660759	BB660759 BB660759
21	280.6	16.5	653	9	BB602405	BB602405 BB602405
22	274.8	16.1	578	9	AV605334	AV605334 AV605334
23	270.8	15.9	628	9	BB660738	BB660738 BB660738
24	269.2	15.8	608	9	AI746617	AI746617 um10d05.Y
25	268.6	15.8	658	9	AI316141	AI316141 u125e04.Y
26	268.2	15.7	901	10	BI763962	BI763962 603049842
27	266.8	15.6	807	10	BI332219	BI332219 602981678
28	265	15.5	492	9	AI663462	AI663462 UK3305.Y
29	260.2	15.3	637	9	BB661317	BB661317 BB661317
30	253.8	14.9	503	10	AA033971	AA033971 z105c09.Y
31	253.4	14.9	565	10	BI143310	BI143310 602907858
32	253	14.8	463	9	AA259000	AA259000 um63c01.Y
33	250.8	14.7	575	9	AA475229	AA475229 um63c01.Y
34	250.8	14.7	630	9	BB661292	BB661292 BB661292
35	247	14.5	915	10	BE788430	BE788430 602114177
36	245	14.4	612	9	BB637428	BB637428 BB637428
37	243.4	14.3	445	9	AA475431	AA475431 um66c08.Y
38	242.8	14.2	490	9	AA106794	AA106794 um33f07.Y
39	238.8	14.0	902	10	BF785293	BF785293 602108493
40	238.6	14.0	612	9	AA611026	AA611026 um75f02.Y
41	237.6	13.9	811	10	BI328477	BI328477 602985750
42	237.6	13.9	607	9	AI956697	AI956697 u17d06.Y
43	236.4	13.9	570	9	AA107600	AA107600 u193a06.Y
44	233.8	13.7	576	9	AI315917	AI315917 u127f05.Y
45	230.8	13.5	451	9	AA012307	AA012307 um07c06.Y

ALIGNMENTS

RESULT 1
AV652088 615 bp mRNA linear EST 15-JAN-2002
AV652088 GLC Homo sapiens cDNA clone GLCWF10 3', mRNA sequence.

AV652088.1 GI:9873102

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 615)
XU.X., HUANG.J., XU.Z., QIAN.B., ZHU.Z., YAN.Q., CAI.T., ZHANG.X.,
XIAO.H., GU.J., LIU.F., HUANG.O., CHENG.Z., LI.N., DU.J., HU.W.,
SHEN.K., LU.G., FU.G., ZHONG.M., XU.S., GU.W., HUANG.W., ZHAO.X.,
HU.G., GU.J., CHEN.Z. and HAN.Z.
Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)

TITLE

JOURNAL
MEDLINE
COMMENT
Contact: Zenguan Han
Chinese National Human Genome Center at Shanghai
351 Guo Shouling Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex. 45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES

source
1..615
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GLCWF10"
/clone.lib="GLC"
/tissue.type="corresponding non cancerous liver tissue"
/dev.stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"

RESULT	3	
AG103175	688 bp	DNA linear GSS 03-NOV-2001
LOCUS	Pan troglodytes DNA, clone: PTB-106108.F, genomic survey sequence.	
DEFINITION	Pan troglodytes DNA, clone: PTB-106108.F, genomic survey sequence.	
ACCESSION	AG103175	
VERSION	AG103175.1 GI:16723692	
KEYWORDS	GSS: GSS (genome survey sequence).	
SOURCE	Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male BAC library, clone:PTB-106108.F.	
ORGANISM	Pan troglodytes	
REFERENCE	Eukaryotes: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi; Mammalia; Eutheria; Primates: Catarrhini: Homiidae; Pan.	
AUTHORS	1 (sites)	
TITLE	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Totoki, Y., Matanabe, H. and Sakaki, Y.	
JOURNAL	BAC end sequences of library PTB	
REFERENCE	Unpublished	
AUTHORS	2 (bases 1 to 688)	
TITLE	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T., Totoki, Y., Matanabe, H. and Sakaki, Y.	
JOURNAL	Direct Submission	
AUTHORS	Submitted (02-AUG-2001) Aaso Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suheiho-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimbesegsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)	
COMMENT	Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the Red process and may have higher chance of clone tracking errors.	
PRIMERS		
Sequencing: -21M13		
LIBRARY		
Vector	: pKS145	
R.Site 1	: SacI	
R.Site 2	: SacI	
Location/Qualifiers		
1.668		
/organism="Pan troglodytes"		
/db_xref="taxon:9598"		
/clone="PTB-106108.F"		
/sex="male"		
/cell_type="lymphoblast"		
/clone_lib="PTB Chimpanzee Male BAC Library"		
BASE COUNT	162 a 160 c 154 g 207 t 5 others	
ORIGIN		
Query Match	22.8%; Score 388; DB 12; Length 688;	
Best Local Similarity	96.4%; Pred. No. 4.3e-95;	
Matches	397; Conservative 0; Mismatches 15; Indels 0; Gaps 0	
177	ttccaatcaatgagccttggagagccttcttcaagtcagttggaagccttggagattca	236
43	ttccggagactagacagctccggagagcctttagtcacaaagttgagagccttggagattca	102
237	gagtcctcatcgtgttcttctcctccctcctcaatgattatcaatcccatatacgtc	296
103	ganccttcatcgtgttcttctcctccctcctcaatgattatcaatcccatatacgtc	162
297	agagaacttctgctgagcagcattccggtcattcgtgctgggtccacatgctgcgaataa	356
163	agagaacttctgctgagcagcattccggtcattcgtgctgggtccacatgctgcgaataa	222
357	tactgagatcgtgaatgaactggaatccctcagtgaagaatgccccttgagaatcctcat	416

Db	223	TACGATCTGTTAANGAAGCTGGAATCCTCAGTGAAGATGCCCTTGAGAAATTTCTAT	282
OY	417	ccactagatcaaatctgaagccagaagaagtgtcgtcgtcttgcattcccccagttgca	476
Db	283	CCGACTAGTACTCAAAATCTGAGGCCACAGAGAAAGTGCTGCTTGCTCATCCCAAGTGCA	342
OY	477	gctcttcacctgaataggagactatccacagacaaagttagggagacacagaacctgtgt	536
Db	343	GCTTCTTCACTGGAAGGAGCTATCCACAGCAACAAGTAGGAGACACAGAAACCTGTCT	402
OY	537	ggatgctgggtatattatgatacaaaagctactcccttgaccattgactgaag	588
Db	403	GGATGGCTGGGTATATGATCAAGACTCTTCCTTGACACATTGTGACTAAG	454
RESULT	4		
AA680184/c		399 bp	EST 19-DEC-1997
LOCUS	AA680184		linear
DEFINITION	z11c10.s1 Soares_fetal_liver_spleen.INFLS.S1 Homo sapiens CDNA		
ACCESSION	AA680184		
VERSION	AA680184.1	GI:2656651	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
JOURNAL	1 (bases 1 to 399)		
COMMENT	Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krisman,D., Kucaba,T., Lacey,M., Le,N., Lennon,G., Mairra,M., Martin, J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,T., Wyllie,T., Waterston,R. and Wilson,R. Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: estewatson.wustl.edu This clone is available royalty-free through LIND ; contact the IMAGE Consortium (Info@image.jnl.gov) for further information. Seq primer: -40m13 fwd. ET from Amersham. Location/Qualifiers 1..399 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="IMAGE:430482" /clone_lib="Soares_fetal_liver_spleen.INFLS.S1" /sex="male" /dev_stage="20 week-post conception fetus" /lib_host="DH10B (ampicillin resistant)" /note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; This is a subtracted version of the original Soares fetal liver spleen INFLS library. 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5', AAGCGAAGAAATTAATTAAGATCTTTTCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."		
BASE COUNT	95 a	100 c	72 g 132 t
ORIGIN			
Query Match	22.7%	Score 387;	DB 9; Length 399;
Best Local Similarity	99.7%	Pred. No. 7.3e-55;	
Matches 398;	Conservative	0; Mismatches	0; Indels 1; Gaps 1.
629	ttcaatcttactctgactgagatgctgtgtggaagacatcatataggtggccatgtccag	688	

```

Db 399 TTCAATTCCTACTTCTGACAGTGTGAGGAGCATATGAGGAGCATGCTCAG 340
689 acggggtgctggtggaatcgtcgtggtggtgtaatacacaataactagatgaagcgt 748
339 ACAGGGGCGCTGGGATCTGCTGGTGTATATCAACATTAATCATGATGAGGCGCT 280
749 taaaggcacttagaagaatgacgcacgaatggaatgaagaatcgtgaagaacccctga 808
279 TAAAGCAGCTTAAAGATGTCACGACCAAAATGGAATTAAGATGCTGAAGAAACCCCTGA 220
809 acatagaagtttgaagatccacatgcaaggaagcgtgcatgcaacagacacaaacta 868
219 ACATGAGGTTGTAAGATCACCATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 160
869 cctgtgtgactgtgtccgcaacccagatgctgtaagaagatcgtatcctgtatctt 928
159 CTGTGTGACTGTTGTCGCCAACCCAGATGCTGTAAGAGATGCTGTAAGATGCTGTAATTT 100
929 tgaag-aaaaaaatcctcaaggaaggaataaataatgattgctacacaaagttacacaa 987
99 TGAGAAAAAAATCTCAAGGAAAAAGCATAAATGATGCTACACAAAGTGAACCA 40
988 tttaagaagcctcatgagcgtgattgtgtgggaattc 1026
39 TTTTAAAGAGCTTCATGAGCTGATTTGTTGGGAAATTC 1

```

```

RESULT 5
LOCUS AO280703 612 bp DNA linear GSS 22-NOV-1998
DEFINITION CITBI-EI-2514B2.TR CITBI-EI Homo sapiens genomic clone 2514B2, DNA
sequence.
ACCESSION AO280703
VERSION AO280703.1 GI:3906522
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

```

```

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 612)
Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and
Venter,J.C.
Use of a random human BAC End sequence Database for Sequence-Ready
Map Building (1998)
Unpublished
Other_GSSs: CITBI-EI-2514B2.TF
Contact: Mark Adams
The Institute of Eukaryotic Genomics
The Institute for Genomic Research
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (inf@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.

```

```

FEATURES
SOURCE Location/Qualifiers
1. 612
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2514B2"
/clone_lib="CITBI-EI"
/sex="male"
/cell_type="sperm"
/notes="Vector: pBelOBAC1, site_1: EcoRI, site_2: EcoRI;
Caltech Human BAC Library D"

```

```

BASE COUNT 242 a 80 c 101 g 189 t
ORIGIN

```

```

Query Match 22.4%; Score 382; DB 12; Length 612;
Best Local Similarity 100.0%; Pred. No. 4.4e-54;
Matches 382; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1324 taaagtcgaaagcgaacaaactactgctacacgacacgaagaatcaaatcaatccaga 1383
Db 1 TAAATTCGAAAGTCAAAACAAATACGTACATGCAACAGAAATCAAAATTAATCCAGA 60
QY 1384 aatagaacctataaataatgcatlaatacgaatgactatccttgaataacagcactga 1443
Db 61 AATGAGACCTATATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 120
QY 1444 aaaggaaggtgataactataataatgactatccttgaataacagcactga 1503
Db 121 AAACGGAAGATTGATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATG 180
QY 1504 tgatgtgacttacttccatggtggaatttaacttcttcttgaataatcttctct 1563
Db 181 TGATGTAAGTACTTTCATGATGATTAATGATTAATGATTAATGATTAATGATTAATGATTA 240
QY 1564 ctgataatttaacaaatagctggtatagttacaaatataaataatgattcaaat 1623
Db 241 CTGTATATTTTAAACAAATAGCTGTATGATTAATGATTAATGATTAATGATTAATGATTA 300
QY 1624 tgaagggcaagggccaggttcaagcaatttcaaacctgatacttcaataaataact 1683
Db 301 TGAAGGCAAAAGGCGGATTCAGCATTTTCAACAGTATGATTAATGATTAATGATTAATG 360
QY 1684 ataattaaaataattatttc 1705
Db 361 ATTAATTAATAAATTAATTAATTC 382

```

```

RESULT 6
LOCUS AO127694 474 bp DNA linear GSS 23-SEP-1998
DEFINITION HS_3091_B2_B02_MF CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate=3091 Col=4 Row=D, DNA sequence.
ACCESSION AO127694
VERSION AO127694.1 GI:3504860
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

```

```

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 474)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3091 Row: D Column: 4
Class: BAC ends
High quality sequence stop: 474.

```

```

FEATURES
SOURCE Location/Qualifiers
1. 474
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=3091 Col=4 Row=D"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/sex="male"
/notes="Organ: sperm; Vector: pBelOBAC1, BAC clones in

```

BASE COUNT 182 a 62 c 70 g 157 t 3 others
 ORIGIN E-Coli DH10B"

Query Match 21.9%; Score 372.8; DB 12; Length 474;
 Best Local Similarity 94.9%; Pred. No. 1.6e-52; Indels 1; Gaps 1;
 Matches 394; Conservative 0; Mismatches 20;

QY 1291 gctgtgtatgaagacatgtgtgaattcattcaaaagttgcaaaagtcacaataact 1350
 Db 4 GGGTTGGGGCTCAGTCACTTANGGCGATTCATAAGTTGCANAGTCANAGCAATCT 63
 QY 1351 gtaacatgcaacgagaataaataaataccagaataagagactataaagcattaa 1410
 Db 64 CTACATGCAACGAAATCAATCAATATCTCTATAGAGACCTATATATATGCAATTT-A 122
 QY 1411 taacatgacttttgacataataagcattgaaacgaaagattagataactaaac 1470
 Db 123 TACATGATCTTTTGACATATATATAGCATTTGAAACGAAAGATTAGATCTAAATAC 182
 QY 1471 attgactatcctttgtaaaacagtcacataagatgtatgactttccatgtgtga 1530
 Db 163 ATGAGCTATCTTTGTAATATGAGTCAATGATGATGATGATGATGATGATGATGAT 242
 QY 1531 attttaaacttttctttgtaatttctctctgtatattttaacaatagctgtga 1590
 Db 243 ATTTTAAATTTCTTTCTTTGTAATTTCTCTGTAATTTTAAACAAATACCTGTA 302
 QY 1591 tagtttaacataataaagatattgtccaattgaaagggcaagggcaggtcagaat 1650
 Db 303 TAGTTTACAAATATTAAGATATTTGTTCAATTTGAAGGGCAAGGCCAGGTGACAAAT 362
 QY 1651 ttccaacgtgtatgacatttaataaataactataaataaataattatattc 1705
 Db 363 TTTCAACGTGTATGATCATTTTAAATAAATACTATTAATAAATAAATAATATATTTC 417

RESULT 7 578 bp DNA linear GSS 22-NOV-1998
 LOCUS AQ280563
 DEFINITION CITBI-EI-2514B1.TR CITBI-EI Homo sapiens genomic clone 2514B1, DNA
 sequence.
 ACCESSION AQ280563
 VERSION AQ280563.1 GI:3906382
 KEYWORDS GSS.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 578)
 AUTHORS Adams M.D., Rounsley S.D., Zhao S., Bass S., Linher K., Golden K.,
 Berry K., Granger D., Suh E., Wible C., Shizuya H., Simon M. and
 Venter J.C.
 TITLE Use of a random human BAC End Sequence Database for Sequence-Ready
 Map Building
 JOURNAL Unpublished (1998)
 COMMENT Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdadams@tigr.org
 Clones are available from Research Genetics (info@resgen.com). BAC
 end search page:
 http://www.tigr.org/tldb/humgen/Bac_end_search/bac_end_search.html.
 Seq primer: M13 Reverse
 Class: BAC ends
 Location/Qualifiers
 1..578
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

BASE COUNT 229 a 79 c 88 g 182 t
 ORIGIN /clone="2514B1"
 /clone_id="CITBI-EI"
 /sex="male"
 /cell_type="sperm"
 /note="Vector: pBelBAC11, Site_1: EcoRI; Site_2: EcoRI;
 Caltech Human BAC Library D"

Query Match 21.7%; Score 370; DB 12; Length 578;
 Best Local Similarity 98.5%; Pred. No. 4.4e-52; Indels 1; Gaps 1;
 Matches 384; Conservative 0; Mismatches 5;

QY 1317 gaattcataagttgcaaaagtcacaataactctgacatgcaacgagaataaataa 1376
 Db 17 GAATTTATTAAGTTGCAAAAGTCAAAACATATCTGATCTGATCTGATCTGATCTGATCT 76
 QY 1377 atccagaataagagactataataatgacatttaatacatgacttttgacataaagc 1436
 Db 77 ATCCAGAAATAGACACTATATTAATGATATGATATGATATGATATGATATGATATGAT 136
 QY 1437 cattgaaacgaaagattagataactaataacatgactatcctcttgtaaacagt 1496
 Db 137 CATTCGAAACGAAAGATGATGATCTTAATATGATGATGATGATGATGATGATGATGAT 196
 QY 1497 caactaagatgtatgactttccatggtgtgaatttaacttttctgtgaatt 1556
 Db 197 CACTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 256
 QY 1557 ttctc-tctgtatatttaacaataagctgtgtatgatttacaataataaagatt 1615
 Db 257 TTTTATATCTGATATTTTAAACAAATGCTGTGTATGATTTTCAATATTAATAAGATAT 316
 QY 1616 gtccaattgaaagggcaagggcaggtcagaatttcaaacgtatgacatttaata 1675
 Db 317 GGTCAATTTGAAGGCCAAGGCCAGGTGACGAATTTTCAACCTGATGATCATTTAATA 376
 QY 1676 aaataactataaataaataattatattc 1705
 Db 377 AAATTAATATTAATTAATAAATAATTAATTTC 406

RESULT 8 367 bp mRNA linear EST 19-DEC-1997
 LOCUS AA682573
 DEFINITION zj20a06.s1 Soares_fetal_liver spleen INFLS-SI Homo sapiens CDNA
 clone IMAGE:450802 3, similar to TR:G1293672 G1293672
 KIDNEY-SPECIFIC TRANSPORT PROTEIN.; mRNA sequence.

ACCESSION AA682573
 VERSION AA682573.1 GI:2659854
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 367)
 AUTHORS Hillier L., Allen M., Bowles L., Dubuque T., Geisler G., Jost S.,
 Krizman D., Kucaba T., Lacy M., Le N., Lennon G., Marra M., Martin
 J., Moore B., Schellberg K., Stepcoe M., Tan F., Theisling B.,
 White Y., Wyllie T., Waterston R. and Wilson R.
 TITLE The Human Genome Project
 JOURNAL Unpublished (1997)
 COMMENT Contact: Wilson R.
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.jnl.gov) for further information.
 Possible reversed clone: similarity on wrong strand
 Seq primer: -40m13 fwd. EF from Amersham

Db	173	TCGTGGAGAAATTTTACAGCGCCATTCTTGTCATCAGTCTGCTGGGGTCCGAACTCTCGA	232
QY	351	caataaactgcagatctgcgtcaatgaaactgcgaatctcagtagaagatgcctcttggaat	410
Db	233	CAATGACACTGATCTCGTAAATAGCATGAGGACCTCGCGCCCCGAGGTCTCTCGAAT	292
QY	411	ctctatccactgactgactcaaatctgaggccagagaagtgctgcttgcttgcactccca	470
Db	293	CTCCATGCCCCCTGAGACCAAACTTGAAAGCAGAAATGTGGGGCTACCTCCACCCCA	352
QY	471	gtggcagctcttccttaactgtagtgagactctcacagcacaaagttaggcagacagaac	530
Db	353	GTGGAGAGCTCTTACCTAGTAATGGAACCTTCCCAATGAGTATGACTATCTGACAGGAGCC	412
QY	531	ctgtgtagatgctggtggtatagatgacaaagctactcctccctgagacttgagtaagt	590
Db	413	CTGTGTGAGAGCGCTGGGGTGTATGACCAAACTCTCTCCACACACTGACTGACTG	472
QY	591	ggacctgtatgtagtattatcagtcagtcaggaatcagtggtgtcaattctactctgactg	650
Db	473	GGACCTCTATGTGATTTATCAGTCTCGAAGAACACAGTGTCCATTTGCATTCATGCTG	532
QY	651	aatgctgtgtaggagcatcataggtggtgcactgctcctagacaggt	694
Db	533	AATGCTGCTGGAGGCTTCTATACGCGCTATCTCTCGACAGGT	576
RESULT	11		
LOCUS	AQ237446/C	741 bp	DNA linear GSS 21-APR-1995
DEFINITION	RPCL11-69K14.TK RPCL11 Homo sapiens genomic clone RPCL11-69K14, DNA sequence.		
ACCESSION	AQ237446		
VERSION	AQ237446.1	GI:3669737	
KEYWORDS	GSS.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 741)		
AUTHORS	Adams M.D., Rounsley S.D., Zhao S., Baas S., Linher K., Golden K., Berry K., Granger D., Suh E., Wible J.C., de Jong P. and Ventre J.C		
TITLE	Use of human BAC End sequences for Sequence-Ready Map Building		
JOURNAL	Unpublished (1998)		
COMMENT	Other-GSSs: RPCL11-69K14.TK		
FEATURES	<p> Contact: Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: mdadams@tigr.org Clones are derived from the human BAC library RPCL11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.bufileo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.bufileo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/cdb/hungen/bac_end_search/bac_end_search.html Seq primer: T7 Class: BAC ends. Location/Qualifiers 1..741 /organism="Homo sapiens" /db_xref="GDB:7526365" /db_xref="taxon:9606" /clone="RPCL11-69K14" /clone_1bp="RPCL11" /sex="Male" /sex="Male" /cell_type="Lymphocytes" /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; RPCL11 Human Male BAC library" 213 a 156 c 188 g 184 t </p>		
BASE COUNT	213 a 156 c 188 g 184 t		
ORIGIN			

Query Match	Best Local Similarity	Score	DB	Length
Matches 387; Conservative	72.3%;	298.2;	10;	820;
		Pred. No. 3.4e-40;	Mismatches 148;	Indels 0; Gaps
160	ctcttgaggagtcacgtgttccacaacaataaggcccttgagagctcttgagaaattga	219		
152	cccatgggagaaagacattcagttccaatggcccttcagagactccgatcaattgga	211		
220	ggcccttgagagattcagatgtgttcacgtgtttatctctccctctcatgtatata	279		
212	agccttgagaaatttcacattccttcagatctctttgtttcttccctgaagccttg	271		
280	atccctcatatactctgtagagaacttgcctgagccatccctgcatcgtgttgctg	339		
272	gtaacctaatattacaattgagaaatttactgcagccattcccaatcattccgctgg	331		
340	cacatgctgagacaatactatcgtatctgtagtaagaacttgaaatccatgaaatg	399		
332	cccatccttggaagacactccctctatattatggcagatgagatcctgagcaagat	391		
400	ctcttgagaatctctatacccaactgactcaaatctgagagccagaaagtgcgtg	459		
392	ctcttgagaaattcctcattcccttgatctcgaactgagaaatgctgctgttt	451		
460	gtccatcccaactgctgagactcttccactgtagatgagactaccagacaaagtga	519		
452	gcccaacacagtgagcgaactcttcttattgaaatggcactttctccaatgagac	511		
520	gacacaaagccgtgtgtagatgctggtgatataatgaagctactctccctgag	579		
512	gacactgagccctgctgtagatggtgggtatgacagagcaactcttctgacatt	571		
580	gtgactaagtggagaccgtgatgtgatatcagtaactgaaatcgaatggttcaat	639		
572	gtgactgagtgaggaccgtgctgtgtaattcagagcactggaattctgttcaatt	631		
640	ctcttgactggaatgctgtgtgtagagcatatagtggtgcatgtctcagaagct	694		
632	ttcatgattggcctattttatgagggcctatcatatggtgacattttgacagac	686		

BASE COUNT 191 a 197 c 185 g 247 t

ORIGIN

/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4972265"
/clone_1db="NCI CGAP Kid14"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: Kidney; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library. "

OY 400 cctttagatcctctccactaggctaataacttagagcgagaagtgtgcgcttt
 | | | | | | | | | | | | | | | | | | | | | |
Db 440 CTCCTAGGATCTCCATCCCCCTGGATTCCAACCTTAGTGAGTAAATGTCTCTTTT
 | | | | | | | | | | | | | | | | | | | | | |
OY 460 gtccatcccagltbgcagctcttcacctgaatggagcatcacagacaagttagga
 | | | | | | | | | | | | | | | | | | | | | |
Db 500 GCCCAACCACAATGGCATTTCTTCATTGAATGGCATTCTTCCAAATGNTCAGGCCA
 | | | | | | | | | | | | | | | | | | | | | |
OY 520 gacacagaacctgttgttatgcttgatgatattgatacaagcttactcccttcagcatt
 | | | | | | | | | | | | | | | | | | | | | |
Db 560 GACTCTGACCCCTGTGTGATGTGGTGTGTATGACAGGACAACCTCTTCTTACATT
 | | | | | | | | | | | | | | | | | | | | | |
OY 580 gtgaactaagtggacctgltatgatattaacgaactaaactagtgltcaattccta
 | | | | | | | | | | | | | | | | | | | | | |
Db 620 GTGACTGATGGGCACTTGTATGTGAATCTCANGCACTAATTCGTCACTAAATATCA
 | | | | | | | | | | | | | | | | | | | | | |
OY 640 ctctcgaactgaatgctgtgggaagcatcataagltggcactgcttcagaacagt
 | | | | | | | | | | | | | | | | | | | | | |
Db 680 TTCAATGATTTGGCCTTNNNTATANGGTATCATATGTGGCAATCTGTGCACACAGGT
 | | | | | | | | | | | | | | | | | | | | | |

[illegible]

REFERENCE	(Bases 1 to 1054)
AUTHORS	NIH-MGC http://mgc.nhl.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

FEATURES	COMMENTS
	<p>Contact: Robert Strassberg, Ph.D. Email: c9abps-r@emil.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM9848 row: d column: 18 High quality sequence stop: 665. Location/Qualifiers</p>

```

SOURCE
1. .1054
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone_image="IMAGE:423793"
/clone_1lb="NCL_CGAP_K1d14"
/lab_host="DH10B (r1 phage-resistant)"
/note="Organ: kidney; Vector: pCMV-SpOr6; Site:1: NCL1;
Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT
Average insert size 1.75 kb. Constructed by Life
Technologies. Note: this is a NCL_CGAP Library. |"
BASE COUNT
276 a 250 c 256 g 272 t

```

Query Match	17.28	Score	292.6	DB	10	Length	1054
Best Local Similarity	73.68	Pred. No.	2.7e-39				
Matches	373	Conservative	0	Mismatches	134	Indels	0
						Gaps	0

QY 188 tggcccttgagagcctctcgatccaagtltgaggccttggagagatlccagatgtcttcac 247
||||| |||| |
Dd 1 TGGCCTTTCAGGAACCTCGAATCAAGTTGGAAGCCGTGGAAGAATTCCAGATCCTTCAGA 60
||||| |||| |

248 tggtttttattctcttcocctctcctatgttatttaaccctacatactgtctagagaacttg 307

[illegible]

Search completed: July 1, 2002, 14:44:55
Job time: 9394 sec

Db 421 CTCAGCCTGTAATTCGTGCTAAATTTTCATTTCATGATTTGCCCTATTATAGGGGCTA 480

QY 668 tcataggtggcatgtctcagacaggt 694

Db 481 TCATATGTGGCCATTTTGTCTAGACAGGT 507

Search completed: July 1, 2002, 14:44:55

Job time: 9394 sec

RESULT 2
US-08-232-463-14/C
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:

QY 1012 ttgttggggaattcgcagaaaaaatcacgagaagaacaccagaagggttttctc 1071
Db 1109 RRATCGAAGTCCCTGCA 1050
QY 1072 cctcaaccacgcagaacacatatttgtacaatatccaa 1113
Db 1049 CCTGCAGCCCAAGCTCGATTATTTCTGTGACCTGAATGGCAA 1008

	Query Match	Best Local Similarity	Score	DB I	Length
Matches	40;	Conservative	228;	Mismatch	194;
				Indels	0;
				Gaps	0;
QY	652 atgctgttgaggagcatatagtgcgatgttcacagagtgctgtgtgaattcgt				711
Db	1469 ATTCGAGTGTGGTTAAAGATAGAGAATTTCGACRRRRRRRRRRRRRRRRR				1410
QY	712 cggtttgtatatccccaataactagaatgagcgcttaagcacctaagaattca	:::::	:::::	:::::	771
Db	1409 RR	:::::	:::::	:::::	1350
QY	772 gcacaatatygataaaatgctgaagaaacctgaacacatagagtgtgaatccacc	:::::	:::::	:::::	831
Db	1349 RR	:::::	:::::	:::::	1290
QY	832 atgcgagagcgctgtagcagcacagaccaaatctactgtgtgaccttccgcaac	:::::	:::::	:::::	891
Db	1289 RR	:::::	:::::	:::::	1250
QY	892 ccagatagtcglaaaagatctgtatcctgtlattttgagaaaaaatctcaaggaaa	:::::	:::::	:::::	951
Db	1229 RR	:::::	:::::	:::::	1170
QY	952 aggagataaaatgttctctacacaaaatgtaaccaatttaagaagctcatyagactga	:::::	:::::	:::::	1011
Db	1169 RR	:::::	:::::	:::::	1110

Query Match	3.2%	Score 54	DB 1	Length 5852
Best Local Similarity	49.1%	Pred. No. 0.00025		
Matches	173	Conservative	0	Mismatches 175; Indels 4; Gaps 1
QY	1345	aatactgtacatgcacccgaatacaataataccgaagaatagagccatataatgc	1404	
Db	5431	AAATTATTTCAATTTTAAAAAATAAAAAAAAAAAAAAAAAATCAATAATATTTTATTG	5490	
QY	1405	atttaatacatgataacttttgacaataagccattggaaacggaagaattgatacta	1464	

Query Match	2.9%	Score 48.8;	DB 2;	Length 1186;
Best Local Similarity	45.7%;	Pred. No. 0.0027;		
Matches 290; Conservative	0;	Mismatches 277;		

QY	1064	ctcttctccctacaaccgacagacatatactagatagactgtaattctcaattataat	11
Db	162	TTTTATTACTTAAATTAATAAAGGAAFTMAAATTAGAAAGCAATATTTAAACAGAAATA	22
QY	1124	ggcataaattgcattatcttaattcaaatctaacttggggacaatgtaactctcttagcaa	11
Db	222	AGCATATGATTCATCTTAAGCGAAACCTAAATAAAGCTGTTCTTGTTAAATCTTGTTGTA	28
QY	1184	tcgtatatttttgggaagtcctctttaaagaattacaattctacaataatctactagta	12
Db	282	AGAAAGGTTCTTCAAAATTAACCTTAACCTAAATGAACCTAAATACGAAACCATACCAATMCA	34
QY	1244	taagatgattccaagaacaaagaanaactcagaattaggatgctggctgctgtagta	130
Db	342	GCACCTGCTATCCAAATAGAGTATCCAGCAACCAATAAATTTGGCAGATTG- AATAA	400
QY	1304	agaccatctgagtgaaattcataagttgcacaaagtcacaacaactactgtaactgaccca	136
Db	401	CATATTATGTAATTAATTATTAATCTTTTAA--GATATTAACCTTTTAAAAAAAAT	457
QY	1364	gaatacaaatatccagaatatagacacctataaactgcatcttaactagatcatt	142
Db	458	AAATGAAATACCTCTATATAAGAAATATATTTTAAATATGTAATAAATGTAATAAATCTTATAT	517
QY	1424	tgcataataagccacttgaaaacggaaggttagctgacataataactgactactct	148
Db	518	AATATTTTAAATAATGATCTGGATTTATACGTTGATCTATTAATCTTAAATAGATGT	577
QY	1484	tgttaataacagtcactaaatgtagttagtcttaactccatggtggaatttaactact	154
Db	578	ATAAGAAATCTTCATATGAAAGAAAGTTTATTTTTCATATCAATCAATAATGTAATCTTAT	637
QY	1544	ttctcttgtaatttctctctgtatatactttaaacaatatagctgtaagtctacaatat	1603
Db	638	TCTATTTTAAATAAATATCGTAATACTTTTAATTAAGAAATA-----AATATTAATAT	689
QY	1604	tataagaatctgtcaaatgaaggcgaaggccaggtcgtaacatcttcaactgcat	1663
Db	690	TTAATAAATAAATAATCTTATTAAGAAGATTTTTCATTTTAATTTTAAAAAGTTAT	749
QY	1664	gtacatttaataaataaactataataaataat	1697
Db	750	ATATCTTAAAAAGATTAATTAATTAATTCATATAT	783

```

1  APPLICATION NUMBER:  US/08/731,722
2  FILING DATE:
3  CLASSIFICATION:  424
4  ATTORNEY/AGENT INFORMATION:
5  NAME:  Whitlock, Ted W.
6  REGISTRATION NUMBER:  36,965
7  REFERENCE/DOCKET NUMBER:  UF-161
8  TELECOMMUNICATION INFORMATION:
9  TELEPHONE:  352-375-8100
10 TELEFAX:  352-372-5800
11 INFORMATION FOR SEQ ID NO:  5:
12 SEQUENCE CHARACTERISTICS:
13     LENGTH:  1186 base pairs
14     TYPE:  nucleic acid
15     STRANDEDNESS:  single
16     TOPOLOGY:  linear
17     MOLECULE TYPE:  DNA (genomic)

```

RESULT 7
 US-08-998-416-541/C
 : Sequence 541, Application US/08998416
 : Patent No. 6239264
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Phillipsen, Peter
 : APPLICANT: Pohlmann, Rainer
 : APPLICANT: Steiner, Sabine
 : APPLICANT: Mohr, Christine
 : APPLICANT: Wendland, Jürgen
 : APPLICANT: Knechtle, Philipp
 : APPLICANT: Redischung, Corinne
 :
 : TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPII
 : TITLE OF INVENTION: AND USBS THEREOF
 :
 : NUMBER OF SEQUENCES: 1152
 :
 : CORRESPONDENCE ADDRESS:
 :
 : ADDRESSEE: No. 6239264artis Corporation
 : STREET: 3054 Cornwalls Road
 : CITY: Research Triangle Park

TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium
 NUMBER OF SEQUENCES: 19
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5369526r1s1
 STREET: One Liberty Place 46th Floor
 CITY: Philadelphia
 STATE: PA
 COUNTRY: USA
 ZIP: 19103
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/867,106
 FILING DATE: 19920625
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: AD PJ 7187
 APPLICATION NUMBER: PCT/AU90/00530
 FILING DATE: 02-NOV-1989
 ATTORNEY/AGENT INFO: AGENT

REGISTRATION NUMBER: 35,134
REFERENCE/DOCKET NUMBER: RICE-0002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 313-487-1111

TELEFAX: 215-568-3439
INFORMATION FOR SEO ID NO: 3

REFERENCE CHARACTERISTICS:
LENGTH: 5852 base pairs
TYPE: null

STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO

NAME/KEY: CDS
LOCATION:

FEATURE:	25/00:0038
NAME/KEY:	CDS

LOCATION: 2378..5038
-07-867-106-2

Query Match

Best Local Similarity	Score
49.78%	44
196	Pred. No.

1192 ttggaatcccttaaaact+

2378 TTTTAAATCTTGATATAAATTAAACCAAACT

1252 ttcagaaacaaagaanaatcacagaaattagatgt

2318 AT-TTACAATAATAATAATAATAATGAAGT

5'cgucygaalccataaaagttgcaaaagtcaaaacaa
| | | | |
2259 AAAAATTTTAAAAAA

1372 ataataatcgaagaattgaa

2199 AATAATATTAATTAAATTTGTCACCAATTTTCTC
||||| | | |

1432 taagccattggaacggaagattaatatctt

2139 T-----TTATAAATAAGCTATAATTAGATAGACAAT

1492 acagtcactaaa--tgatgttagttactttccatg

...CAAAACCAAGTAATTAATTTATTAATC

1

Page 6


```

? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/107,755A
? FILING DATE: 19-AUG-1993
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/827,658
? FILING DATE: 30-JAN-1992
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/657,584
? FILING DATE: 19-FEB-1991
? ATTORNEY/AGENT INFORMATION:
? NAME: Salivanchuk, David R.
? REGISTRATION NUMBER: 31,794
? REFERENCE/DOCKET NUMBER: UF114.C2
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (904) 375-8100
? TELEFAX: (904) 375-8100
? INFORMATION FOR SEQ ID NO: 8:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1511 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: unknown
? MOLECULE TYPE: DNA (genomic)
? ORIGINAL SOURCE:
? ORGANISM: Amsacta moorei entomopoxvirus
? FEATURE:
? NAME/KEY: CDS
? LOCATION: complement (18..218)
? FEATURE:
? NAME/KEY: CDS
? LOCATION: complement (234..782)
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 852..1511
US-08-107-755A-8

Query Match 2.6%; Score 43.6; DB 1; Length 1511;
Best Local Similarity 45.5%; Pred. No. 0.063;
Matches 282; Conservative 0; Mismatches 324; Indels 14; Gaps

QY 1076 caaccagcaagaacatattatgatatacagaacatcacaattataatgatgcaattatg 1135
Db 825 CAATGAAAAAATAAATAATATCAAAATGATTTACTAATTCATATTAATTTAATA 884
QY 1136 cattattccaataactgttg9gg9gcafgtaactcctcgagcaatcgtatatltt 1195
Db 885 AATATTTTAAATATATATATTTAAATAAATAATATATATTAACAGACATATTTTAAAT 944
QY 1196 g9g9agtcctttaaagaagttacaattatcaataattcactagtagaagaagatcca 1255
Db 945 ATTAATATATATTAATAAATAATAGTTATTTAGAGAAGATTCATATATATATATATAT 1004
QY 1256 gaacacaaagaacatcacagaattggatggtgctg9gctg9t9tgaagcacatgta 1315
Db 1005 AATATTTTAAATATATATTTCCAGAAATATTTAAAGTTTATATATTTCAA-----AT 1055
QY 1316 tgaattcataaagttgcasaagtcasaacaactactgtacatgaaccgagaatccaata 1375
Db 1056 TTAATATATTTAATTTTAAATTTTATACAAATTTAAAAAATATATACATATTTAGATATA 1115
QY 1376 aatccagaagaatagagacctatataatgcatcttatacatgatactctttgacataaag 1435
Db 1116 TCTTATTAACAAAAATAGCAATATTAAGTAATATTTACTACACACTTCTATAGAAATTTT 1175
QY 1436 ccattggaacacggaagatgagatagactaataataacattgctatcctt9tgaataacg 1495
Db 1176 AATGTGATCATTTATATATTAATGACTATA---ATTATTTAATAAATTTAGTAATAATT 1231
QY 1496 tcaataaatgatgttagttactcttccatg9t9ggaattcaattactactcttcttgtaat 1555
Db 1232 AAAAAAATTTATATATATCTAAAAATAAATTGGTACTACTTTAAATAATAGTTTTCCTATTTAG 1291

```

OY 1556 ttctctcgtatatttaacaatagctgtagttcaaatatataaagatt 1615
 Db 1292 TATAGTTGAGTTAAATATGCAATCAATAAATAAGATTAATATATGAAAAATT 1351
 OY 1616 gtcaaatggaagcgaag-gccaggttcagcaatttcaacgctatgtaatt 1674
 Db 1352 AATTAATTTAAAAAATTAAGATATCTTTCAATGTTAAAAAATAATATACATTGGAT 1411
 OY 1675 aaataactataaataa 1694
 Db 1412 AAAATTTCCAAAAAGTATA 1431

RESULT 14

US-08-544-332-8
 ; Sequence 8, Application US/08544332
 ; Patent No. 5935777

GENERAL INFORMATION:

APPLICANT: Moyer, Richard W.
 APPLICANT: Hall, Richard L.
 APPLICANT: Gruidl, Michael E.
 TITLE OF INVENTION: No. 5935777e1 Entomopoxvirus Expression System
 NUMBER OF SEQUENCES: 77
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Gerard H. Bencen
 STREET: 2421 N.W. 41st Street, Suite A-1
 CITY: Gainesville
 STATE: FL
 COUNTRY: USA

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/544,332
 FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/991,867
 FILING DATE: 07-DEC-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/107,755
 FILING DATE: 19-AUG-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: WO 92/14818
 FILING DATE: 12-FEB-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/827,685
 FILING DATE: 30-JAN-1992

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US 07/557,584
 FILING DATE: 19-FEB-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Bencen, Gerard H.
 REGISTRATION NUMBER: 35,746
 REFERENCE/DOCKET NUMBER: DF114.C4
 TELEPHONE: 904-375-8100
 TELEFAX: 904-372-5800

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1511 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: unknown
 MOLECULE TYPE: DNA (genomic)
 ORIGINAL SOURCE:
 ORGANISM: Amsacta moorei entemopoxvirus
 FEATURE:
 NAME/KEY: CDS

LOCATION: complement (18..218)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: complement (234..782)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 852..1511
 US-08-544-332-8

Query Match 2.6%; Score 43.6; DB 2; Length 1511;
 Best Local Similarity 45.5%; Pred No. 0.063;
 Matches 282; Conservative 0; Mismatches 324; Indels 14; Gaps 3;

OY 1076 caaccagcaagacatattatgatacgaatcctaataatataatgcatattg 1135
 Db 825 CATGAAAAAATAAATAATATCAAAATGATTTACATAATTCGATATATTTAA 884
 OY 1136 cattattcaaatctggtggtgacatgtaattctctgagcaatctgatat 1195
 Db 885 AATATTTAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 944
 OY 1196 gggaggtccttaaaagttcaaatattacataattactagtagatgattca 1255
 Db 945 ATATATATATTAATAAATAATATGTTATAGAGAAATTCGATATATTAAT 1004
 OY 1256 gaacaaaagaataatcaagattagatgctgctgctgctgctgctgctgca 1315
 Db 1005 AATATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1055
 OY 1316 tgaattcaaaagttcaaaagttcaaaagttcaaaagttcaaaagttcaaa 1375
 Db 1056 TTAATATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1115
 OY 1376 aatcagaataagagactataataatgacttaataatgacttaataatg 1435
 Db 1116 TCTTATTAACAAATAATACCAATAATTAATTAATTAATTAATTAATTA 1175
 OY 1436 ccatggaaaacggaagattagataactaaataactgactgactctctgt 1495
 Db 1176 AATGTGAATCATGTAATTAATTAATTAATTAATTAATTAATTAATTA 1231
 OY 1496 tcaataatgattgtagtctcttccatggtggaatttaattacttcttct 1555
 Db 1232 AAAAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1291
 OY 1556 ttctctcgtatatttaacaatagctgtagttcaaatatataaagatt 1615
 Db 1292 TATAGTTGAGTTAAATATGCAATCAATAAATAAGATTAATATATGAAAAATT 1351
 OY 1616 gtcaaatggaagcgaag-gccaggttcagcaatttcaacgctatgtaatt 1674
 Db 1352 AATTAATTTAAAAAATTAAGATATCTTTCAATGTTAAAAAATAATATACATTGGAT 1411
 OY 1675 aaataactataaataa 1694
 Db 1412 AAAATTTCCAAAAAGTATA 1431

RESULT 15

US-08-998-416-1137/c
 ; Sequence 1137, Application US/08998416
 ; Patent No. 6239264

GENERAL INFORMATION:

APPLICANT: Phillippen, Peter
 APPLICANT: Pohlmann, Rainer
 APPLICANT: Steiner, Sabine
 APPLICANT: Mohr, Christine
 APPLICANT: Wendland, Jurgen
 APPLICANT: Knechtle, Philipp
 APPLICANT: Reibschung, Corinne
 TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
 TITLE OF INVENTION: AND USES THEREOF

Tue Jul 2 10:13:51 2002

us-09-674-235-19.rni

Page 10

NUMBER OF SEQUENCES: 1152
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6239264artis Corporation
STREET: 3054 Cornwallis Road
CITY: Research Triangle Park
STATE: No. 6239264th Carolina
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/998,416
FILING DATE: 24-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: CH 0016/97
FILING DATE: 31-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Meligs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 1137:
SEQUENCE CHARACTERISTICS:
LENGTH: 636 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: PAGI692RP
US-08-998-416-1137

Query Match	Score	DB	Length
2.58;	43.4;	4;	636;

Best Local Similarity 44.28; Pred. No. 0.05;
Matches 223; Conservative 0; Mismatches 281; Indels 1; Gaps 1

[illegible]

Db 186 ATATTATATATAATTAAATGATGAATTAGCTAAATTATATTAATTAATTAATAA 127

Qy 1680 aactataaatataaatatattt 1704

Db 126 GTATTAAATCAATCAATTAATTAATTT 102

Search completed: July 1, 2002, 15:40:48
Job time: 10497 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

```
Run on:      July 1, 2002, 15:45:51 ; Search time 316.64 seconds
              (without alignments)
              9245.007 Million cell updates/sec
```

Title: US-09-674-235-19
1705

Sequence: 1 aagaactgaggaagctctt.....aaattaaaaattatattc 1/00

Scoring table:

	IDENTITY_NUC
Canoo	10.0
Ganext	1.0

Total number of hits satisfying chosen parameters: 34728/4

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match	0%
	Maximum Match	100%

1: N_Geneseq_03.8802.*
2: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT.*
3: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT.*
4: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT.*
5: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT.*
6: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT.*
7: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT.*
8: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT.*
9: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT.*
10: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT.*
11: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT.*
16: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT.*
17: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT.*
18: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT.*
19: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT.*
20: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT.*
21: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT.*
22: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001C.DAT.*
23: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT.*
24: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1705	100.0	1705	21	AA238317	Human transmembran
2	1000	58.7	1030	22	AAH98847	Human EST-derived
3	804	47.2	804	21	AA238316	Human transmembran
4	801.8	47.0	2684	22	AAK85824	HOAT5 DNA. Homo s
5	773	45.3	783	22	AAK70136	Human immune/Baema
6	761	44.6	782	22	AAK70138	Human immune/Baema
7	759.4	44.5	782	22	AAK70137	Human immune/Baema
8	738.6	43.3	1316	22	AAK55043	Nucleotide sequenc
9	687	40.3	1353	22	ABA08976	Human secreted pti

10	399	23.4	1375	22	AA1599315	Human poly(ADP-ribose)
11	399	23.4	1986	22	AA059631	Human transporter
12	395.8	23.2	2027	21	AA6168923	CDNA encoding a hu
13	378	22.2	1977	22	AA685823	hOAT4 DNA. Homo s
14	232.8	13.7	1132	22	AA161001	Human poly(ADP-ribose)
15	218.4	12.8	2865	22	AA612566	Human transporter
16	208.8	12.2	2455	22	AA017476	Human transporter
17	180.8	10.6	2047	22	AA017480	Human protein havi
18	180.8	10.6	2194	22	AA012607	Human organic anio
19	180.8	10.6	2210	22	AA032614	Human foetal liver
20	161.8	9.5	540	22	AB03164	Probe #8860 for ge
21	161.8	9.5	540	22	AB030394	Human brain expres
22	161.8	9.5	540	22	AAK11579	Human brain expres
23	161.8	9.5	540	22	AAK37346	Human bone marrow
24	161.8	9.5	540	22	AA118114	Probe #1817 for ge
25	161.8	9.5	540	22	AA143191	Probe #11877 used
26	145.2	8.5	447	22	AA107074	Human reproductive
27	145.2	8.5	9370	22	AA107075	Human foetal liver
28	136.2	8.0	339	22	AB055594	Probe #18672 for g
29	136.2	8.0	339	22	AB040126	Human brain expres
30	136.2	8.0	339	22	AAK24178	Human bone marrow
31	136.2	8.0	339	22	AAK50227	Probe #17241 for g
32	136.2	8.0	339	22	AA127308	Probe #24865 used
33	136.2	8.0	339	22	AA156179	Rat EST-derived co
34	134.4	7.9	685	22	AAH98991	Human CDNA 5'-end
35	108	6.3	789	22	AAK92269	Human CDNA clone r
36	108	6.3	789	22	AAK93561	Human full-length
37	108	6.3	1871	22	AAK94561	Rat liver toxicolo
38	98.8	5.8	280	21	AA010196	Human reproductive
39	98.2	5.8	419	22	AA102433	Rat organic anion
40	88.6	5.2	2294	20	AAV79584	Human poly(ADP-ribose)
41	85.2	5.0	379	22	AA183356	Human organic anion
42	83.8	4.9	2123	21	AAZ29300	Human organic anion
43	83.8	4.9	2127	22	AAAC85819	hOAT1 DNA. Homo s
44	83.8	4.9	2171	20	AAV79585	Human organic anion
45	81	4.8	1638	22	AAAC8044	Mouse organic anion

ALIGNMENTS

RESULT	1
ID	AAZ38317 standard; cDNA; 1705 BP.
XX	
AC	AAZ38317;
DT	09-FEB-2000 (first entry)
XX	
DE	Human transmembrane protein cDNA clone HP02000.
XX	
XX	HP02000; transmembrane domain; liver; expression; homology;
XX	organic cation transporter; drug excretion; antibody; assay reagent;
KW	diagnostic marker; primer; probe; antisense; gene therapy;
KW	agonist; antagonist; ligand; therapeutic; ds.
XX	
OS	Homo sapiens.
XX	
XX	
FH	Key
FT	Location/Qualifiers
FT	187..993
FT	/*tag= a
FT	/product= "Human transmembrane protein HP02000"
XX	
PN	W09955862-A2.
XX	
PD	04-NOV-1999.
XX	
PF	27-APR-1999; 99WO-JP022226.
XX	
PR	28-APR-1998; 98JP-0119395.
XX	
XX	
PA	(SAGA) SAGAMI CHEM RES CERT.
PA	(PROT-) PROTEGENE INC.

QY 1681 actataaataaataatatttc 1705
 |||
 DB 1681 actataaataaataatatttc 1705

AAH98847
 ID AAH98847 standard; cDNA; 1030 BP.

AAH98847;

12-OCT-2001 (first entry)

Human EST-derived coding sequence SEQ ID NO: 704.

Human: sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
 tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
 diagnostics; forensic test; gene mapping; genetic disorder;
 biodiversity; gene therapy; nutrition; ss.

Homo sapiens.

W0200154477-A2.

02-AUG-2001.

25-JAN-2001; 2001WO-US02687.

25-JAN-2001; 2000US-0491404.

17-JUL-2000; 2000US-0617746.

03-AUG-2000; 2000US-0631451.

15-SEP-2000; 2000US-0663870.

(HYSE-) HYSEQ INC.

Tang XT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
 Cao Y, Dermanac RA, Zhang J, Werhman T;

WPI: 2001-476164/51.

P-PSDB; AAM24188.

Isolated polypeptide for treatment of diseases, diagnostics, raising
 antibodies and research use -

Claim 1; Page 638; 1275pp; English.

The present invention provides the protein and coding sequences of novel
 proteins from a variety of organisms, including human, dog, cat, horse,
 cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
 urchin and tomato. These were derived from expressed sequence tags (ESTs)
 from the organism of interest. They can be used in diagnostics,
 forensics, gene mapping, identification of mutations, to assess
 biodiversity and for nutritional purposes. The present sequence is a cDNA
 of the invention.

Sequence 1030 BP; 289 A; 217 C; 247 G; 277 T; 0 other:

Query Match 58.7%; Score 1000; DB 22; Length 1030;
 Best Local Similarity 99.8%; Pred. No. 1.7e-218;
 Matches 1022; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 8 gaggaagctcttcactacgctgtatgcaactgtgagtcggcgccatgagatgaga 67
 |||
 DB 7 gaggaagctcttcactacgctgtatgcaactgtgagtcggcgccatgagatgaga 66
 |||
 QY 68 attatgagggatcaataaagcttaattgaaataaagaagaataatttcctcct 127
 |||
 DB 67 attatgagggatcaataaagcttaattgaaataaagaagaataatttcctcct 126
 |||
 QY 128 ttgaactatctcgttaagcattgtgctcctctggtgggtgacgtgttcacatcaa 187
 |||
 DB 127 ttgaactatctcgttaagcattgtgctcctctggtgggtgacgtgttcacatcaa 186

QY 188 tggccttggagagctcttgaagcttggagagccttggagagattcagatgtctcacc 247
 |||
 DB 187 tggccttggagagctcttgaagcttggagagccttggagagattcagatgtctcacc 246
 |||
 QY 248 tggctttatctctcctctctcctatgtattatccctcatatctgtagaacttgg 307
 |||
 DB 247 tggctttatctctcctctctcctatgtattatccctcatatctgtagaacttgg 306
 |||
 QY 308 ctgcagcattccctgctcatcgttgcctgggtccacatgctggacataactgactg 367
 |||
 DB 307 ctgcagcattccctgctcatcgttgcctgggtccacatgctggacataactgactg 366
 |||
 QY 368 gtaatgaactggaaactcctcagtgaaagtgcctcttgagatctctatccactagact 427
 |||
 DB 367 gtaatgaactggaaactcctcagtgaaagtgcctcttgagatctctatccactagact 426
 |||
 QY 428 caaatctgagcagagagatgctgccttgccttccatccctgagcttctcacc 487
 |||
 DB 427 caaatctgagcagagagatgctgccttgccttccatccctgagcttctcacc 486
 |||
 QY 488 tgaatggactatccacagacaaagtgagagacacagacagacagacagacagacag 547
 |||
 DB 487 tgaatggactatccacagacaaagtgagagacacagacagacagacagacagacag 546
 |||
 QY 548 tatatgataaagcttactcctctgacacattgtgactaagtggagacctgtatgtatt 607
 |||
 DB 547 tatatgataaagcttactcctctgacacattgtgactaagtggagacctgtatgtatt 606
 |||
 QY 608 atcagcactgaaatcagtgcttcaattcctacttctgactggaatgctgtgtgagagca 667
 |||
 DB 607 atcagcactgaaatcagtgcttcaattcctacttctgactggaatgctgtgtgagagca 666
 |||
 QY 668 tcataggtggcactgtctcagacaggtgctgtgtggaatctgtgtgtgataatca 727
 |||
 DB 667 tcataggtggcactgtctcagacaggtgctgtgtggaatctgtgtgtgataatca 726
 |||
 QY 728 ccataaactagatgaggtcttaagagcacttgaagaagtgcagacacaaatggataa 787
 |||
 DB 727 ccataaactagatgaggtcttaagagcacttgaagaagtgcagacacaaatggataa 786
 |||
 QY 788 agaatgctg-aagaaacctggaacatagaggtgtgtaagatccacacacagagagctg 846
 |||
 DB 787 agaatgctg-aagaaacctggaacatagaggtgtgtaagatccacacacagagagctg 846
 |||
 QY 847 gatcagacagacaaactactctgtgtgactgtgttcgcgaacccagatgctgtaa 906
 |||
 DB 847 gatcagacagacaaactactctgtgtgactgtgttcgcgaacccagatgctgtaa 906
 |||
 QY 907 aggatctgatactgtgtaattttgag-aaaaaatctaaaggaagaagcataaaatga 965
 |||
 DB 907 aggatctgatactgtgtaattttgagaaaaaaatctaaaggaagaagcataaaatga 966
 |||
 QY 966 ttgtcaccaaaagtgacaaattttaagaagccttcatagtgctgattgtgtggaaatt 1025
 |||
 DB 967 ttgtcaccaaaagtgacaaattttaagaagccttcatagtgctgattgtgtggaaatt 1026
 |||
 QY 1026 caga 1029
 |||
 DB 1027 caga 1030

RESULT 3
 AAZ38316
 ID AAZ38316 standard; cDNA; 804 BP.
 XX
 AC AAZ38316;
 DT 09-FEB-2000 (first entry)
 XX Human transmembrane protein cDNA clone HP02000 coding sequence.
 DE Human transmembrane domain; liver; expression; homology;
 XX HP02000; transmembrane domain; liver; expression; homology;


```
Db 1814 gaaagagcacaataatgattgtctacacaaagcacaattttaagaagccttcacag 1873
Qy 1008 ctgattgttgagggaatttcagaaaaaaatacaggaagaaacacccagagggttt 1067
Db 1874 ctgattgttgagggaatttcagaaaaaaatacaggaagaaacacccagagggttt 1067
Qy 1068 ttctctacacacagcagaacacacacacacacacacacacacacacacacacacac 1333
Db 1934 ttctctacacacagcagaacacacacacacacacacacacacacacacacacacac 1333
Qy 1128 ttaatttgattttttttaaaatttaactctgtggagacatgtaactctctgagcaatcg 1187
Db 1994 ttaatttgattttttttaaaatttaactctgtggagacatgtaactctctgagcaatcg 2053
Qy 1188 atatttttgggaagtcctttaaaagttaacaaatttaacaaatttaacaaatttaacaa 1247
Db 2054 atatttttgggaagtcctttaaaagttaacaaatttaacaaatttaacaaatttaacaa 2113
Qy 1248 atgtttcagaacacacacacacacacacacacacacacacacacacacacacacac 1306
Db 2114 atgtttcagaacacacacacacacacacacacacacacacacacacacacacacac 2173
Qy 1307 accatgtgatgaattcacaagttgcagaagtcacaacacacacacacacacacacacac 2173
Db 2174 accatgtgatgaattcacaagttgcagaagtcacaacacacacacacacacacacacac 2233
Qy 1367 atcaaaataatccagaacacacacacacacacacacacacacacacacacacacacac 1426
Db 2234 atcaaaataatccagaacacacacacacacacacacacacacacacacacacacacac 2293
Qy 1427 cataataagcattggaacacacacacacacacacacacacacacacacacacacacac 1486
Db 2294 cataataagcattggaacacacacacacacacacacacacacacacacacacacacac 2353
Qy 1487 taataacagtcactaaatgtatgtactcttcacacacacacacacacacacacacac 1547
Db 2354 taataacagtcactaaatgtatgtactcttcacacacacacacacacacacacacac 2413
Qy 1547 ctgttaattttctctctgtatattttaacacacacacacacacacacacacacacac 1606
Db 2414 ctgttaattttctctctgtatattttaacacacacacacacacacacacacacacac 2473
Qy 1607 aagatattgttcaaatgaagggcagaagcaggttcagcaatttcaacacacacacac 1666
Db 2474 aagatattgttcaaatgaagggcagaagcaggttcagcaatttcaacacacacacac 2533
Qy 1667 catthaataataataataataataataataataataataataataataataataata 1705
Db 2534 catthaataataataataataataataataataataataataataataataataata 2572

RESULT 5
AAK70136/C
ID AAK70136 standard; DNA; 783 BP.
AC AAK70136;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:24948.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX
KW cytosolic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN W0200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001MO-US01354.
XX
```

```
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-MAR-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214686.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 25-SEP-2000; 2000US-0234999.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
```


PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239353.
 PR 13-OCT-2000; 2000US-0239367.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-024617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246539.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251858.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251899.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI: 2001-483426/52.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -

XX
 PS Disclosure; SEQ ID NO 24948; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 783 BP; 242 A; 119 C; 105 G; 317 T; 0 other;

Query Match 45.3%; Score 773; DB 22; Length 783;
 Best Local Similarity 100.0%; Pred. No. 9.1e-167;
 Matches 773; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 933 aaaaaaaatcagaagaaagcagataaaatgattctcacaaaagtaccaaatttta 992
 Db AAAAAAAAAATCTCAAGGAAAGGCAATMAAAATGTTGCTACACAAAAGTGCAAAATTTTA 717
 QY 993 agaagcctcatagcctgattggtggygaaatcagaagaaataaacaggaagaaaga 1052
 Db AGAAGCCTTCATGAGTGTGATGGGGAATTCAGAAAAAAATACAGAAAGACAA 657
 QY 1053 caccagaaggggtttttccctacacacagcaagaaacatatagatatacatca 1112
 Db CACCAAGAGGGGTTTTTCCCTACACACAGCAAGCAACATATATATGATCATGAATCTCA 597
 QY 1113 attataatgatgcatiaatttgcatltaattcaaaataacttgyggggagcatgta 1172
 Db ATTATTAATTAATGCAATTAATTTGCAATTTATTCATAAATTAATCTGGGACATGTAAT 537
 QY 536 TTTACTAGTAGATTAATGATGATTCAGAAACAAAAGAAATCAGAAATTAATGATGCTGG 417
 QY 1173 ctcttgagcaatcgtatatttttgggaagtcctttaaagaattcaaatatataca 1232
 Db CTCTTGAGCAATCTGATATATTTTGGGAAGTCCITTTAAAAATTTCAAATTAATCAATAAA 477
 QY 1233 ttactagtagataagaatgattcagaagaaagaaatcacagaattagatgtgctgtg 1292
 Db TTTACTAGTAGATTAATGATGATTCAGAAACAAAAGAAATCAGAAATTAATGATGCTGG 417
 QY 476 TTTACTAGTAGATTAATGATGATTCAGAAACAAAAGAAATCAGAAATTAATGATGCTGG 417
 QY 1293 ctgtgtatgaagcaacatgtgataatcataaagttgcaaaagttcaaaacaactgt 1352
 Db CTGTGTATGAGCAACATGATGATTAATTCATTAAGTTGCAAAATCAAAACAAATACGT 357
 QY 416 CTGTGTATGAGCAACATGATGATTAATTCATTAAGTTGCAAAATCAAAACAAATACGT 357
 Db CTGTGTATGAGCAACATGATGATTAATTCATTAAGTTGCAAAATCAAAACAAATACGT 357
 QY 1353 acatgcaacagaagaatcaaaataaataacagaagacatataaaatgacattta 1412
 Db ACATGCAACGAGAAATCAAAATTAATTAATGAGACCTATTAATGATGATTAATTA 297
 QY 356 ACATGCAACGAGAAATCAAAATTAATTAATGAGACCTATTAATGATGATTAATTA 297
 QY 1413 catgatactttgacataaagcaatgtgaaacaggaagttgataacttaataa 1472
 Db CATGATACCTTTTGACATTAATTAAGCCATTGGAAACGGAAGATTTGATATCTTAATTAACAT 237
 QY 236 CATGATACCTTTTGACATTAATTAAGCCATTGGAAACGGAAGATTTGATATCTTAATTAACAT 237
 Db CATGATACCTTTTGACATTAATTAAGCCATTGGAAACGGAAGATTTGATATCTTAATTAACAT 237
 QY 1473 tgaactatccttgtataaagtcactaataatgattgattactttccatggtggaat 1532
 Db TGACTATCTCTTGTAAATACAGTCACTAAATGATGTTAGTTACTTTCCATGAGTGGAAAT 177
 QY 1533 ttaataacttttcttctgtaattttctctgtataltttaacaatagctgtgata 1592
 Db TTTAATTAATCTTTCTTGTAAATTTTCTCTGTGATTAATTTAAACAAATGAGCTGATATA 117
 QY 1593 gttaacaatatataagaatattgttcaaatgtgaaggcaagggcaggtttcagcaattt 1652

Db 116 GTTTAAATATATTAAGATATGTCATTAATTAAGCGCAAGCCAGGTCAGCAATTT 57
Qy 1653 tcaactgtatgtacactttaataactaataaataataataataatatttc 1705
Db 56 TCAAACTGTATGTACATTTAAATAAATACTAATAATTAATAATATATTTTC 4
RESULT 6
AAK70138/c
ID AAK70138 standard; DNA; 782 BP.
XX
AC AAK70138;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:24950.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX
KW cyclostatic; gene therapy; vaccine; metastasis; ds.
OS Homo sapiens.
XX
PN W0200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001MO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0190076.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-020515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 26-JUL-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.

PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249289.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 06-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PR
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 DR WPI: 2001-483426/52.
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 PR
 PS Disclosure: SEQ ID NO 24950; 3071bp + Sequence Listing: English.
 XX
 XX
 CC AAK4951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytosolic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 CC
 CC
 SQ Sequence 782 BP; 241 A; 119 C; 105 G; 317 T; 0 other;

Query Match 44.6%; Score 761; DB 22; Length 782;
 Best Local Similarity 99.9%; Pred. No. 5e-164; 0; Indels 1; Gaps 1;
 Matches 772; Conservative 0; Mismatches

QY 933 aaaaaaatctcaagaaagcgtataaagatgctgtacacaaagtgcacaaatttca 992
 DB 775 AAAAAAAATCTCAGAGAAAGGCATTAATGATTCCTACACAAAGACCAATTTTA 716
 QY 993 agagcctcatgagctgattggtgggaaatcagaaaaaaatcagagaagaaca 1052
 DB 715 AGAGCCTCATGAGCTGATGTGTGGGGAATTCAGAAAAAAATACAGAAAAAGACA 656
 QY 1053 caccagaaggggttttttcctcaaacccagcaagacatatactagatcagaaatc 1112
 DB 655 CACCAGAGGGGTTTTCCTCAACACGACAGAAACATATATGATACATGAATTCACA 596
 QY 1113 attataatgagcatattgcatcttattcacaataactggtgggagcatgata 1172
 DB 595 ATTATATATGAGCATTAATTTGCAATTTATTTCAAAATTAACCTGTGGGACATGTAA 536

QY 1173 ccttgagcaatctgatatatttgggaagtccttaaaaaagttacaattatcaataa 1232
 DB 535 CCTTGAGCAATCTGATATTTTGGGAAGTCCTTAATAAGTTACAAATTTATCAATAA 476
 QY 1233 ttactagatagaatgattcagaacaaagaanaatcagaagatagagtgctg 1292
 DB 475 TTACTAGTATGATATGATGATTCAGAAACAAAGAAATCACAGAAATTAAGATGCGCTGG 416
 QY 1293 ctgggtatgagagcccatggtatgaattcaataagttgcagaagtcacaataatct 1352
 DB 415 CTGGGTATGAGCCACCATGTCATGATTAATCATTAAGTTGCAAAAGTCAAAACATACTGT 356
 QY 1353 acatgaaccagaatcacaataaataccagaataatgagagccatataatgattata 1412
 DB 355 ACATGCAACCCAGAAATCAAAATTAATCCAGAAATGAGACCTATTAATGATTAATA 296
 QY 1413 catgatactttgaactaataagccattggaagcgaagatgagataactaataaact 1472
 DB 295 CATGATCTTTGACATTAATTAAGCCATGTGAAACGAAAGATTAGATCTTAATPACT 236
 QY 1473 tgactatccttggtaataacagtcactaataatgattgattgaactcttcacatggtgaat 1532
 DB 235 TGACTATCTCTTGTAAATACAGTCACTAAATGATGTTAGTTACTTTTCATGTGTAAT 176
 QY 1533 tttaactctttcttctgtaatttctctctgtatatttaacaataagctgata 1592
 DB 175 TTTAATTAATCTTTCTTTCTTGTAA-TTTTCTCTGTATATTTTAACAAATGCTGATTA 117
 QY 1593 gttacaatatataagatatcttcaaatggaagcgaagccagttcagcaatt 1652
 DB 116 GTTCAATATATTAATAAGATTTGTTCAATTTGAAGGCGCAAGCCAGGTTCAAGCAATT 57
 QY 1653 tcaaatgattgacatttaataaataactaataaataaattatattc 1705
 DB 56 TCAAACTGATGATCTTAATTAATAATTAATAATTAATAATTAATTTTC 4
 RESULT 7
 ID AAK70137 standard; DNA; 782 BP.
 XX AAK70137;
 AC
 XX 06-NOV-2001 (first entry)
 DT
 XX Human immune/hematopoietic antigen genomic sequence SEQ ID NO:24949.
 DE Human: immune; haematopoietic; immune/hematopoietic antigen; cancer;
 KW Human: immune; haematopoietic; immune/hematopoietic antigen; cancer;
 KW cytosolic; gene therapy; vaccine; metastasis; ds.
 XX
 OS Homo sapiens.
 XX
 XX WO200157182-A2.
 PM
 XX 09-AUG-2001.
 PD
 XX 17-JAN-2001; 2001WO-US01354.
 PF
 XX 31-JAN-2000; 2000US-0179055.
 XX 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-MAR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205451.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 14-AUG-2000; 2000US-0220968.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225477.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 14-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226682.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 05-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 06-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 08-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0232081.
 PR 14-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 21-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 25-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 26-SEP-2000; 2000US-0235498.
 PR 27-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 29-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250160.
 PR 05-DEC-2000; 2000US-0250393.
 PR 05-DEC-2000; 2000US-0251988.
 PR 06-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 (HUMA-) HUMAN GENOME SCI INC.
 PA Rosen CA, Barash SC, Ruben SM;
 PI WPI; 2001-483426/52.
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 XX useful for preventing, diagnosing and/or treating cancers and
 XX metastasis -
 PS Disclosure; SEQ ID NO 24949; 3071bp + Sequence Listing; English.
 XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. (I)
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I) by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,

CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK5442 to AAK4950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 CC
 XX Sequence 782 BP; 242 A; 119 C; 105 G; 316 T; 0 other;

Query Match 44.5%; Score 759.4; DB 22; Length 782;
 Best Local Similarity 99.7%; Pred. No. 1.1e-163;
 Matches 771; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

QY 933 aaaaaaacctcaagaaagcacaataatgtctacacaaagtgcacaatttta 992
   |||||||
DB 775 AAAAAAACTCAAGAAAGGCAATAAATGCTCTACACAAAAGTGCACAAATTTTA 716
   |||||||
QY 993 agaagcctcatgagctgtgtggtggaatcagaataaataacaggaagaaca 1052
   |||||||
DB 715 AGAAGCCTTCATGAGCTGATGTGGGGAATTCAGAAAAAAATACAGAAAGACAA 656
   |||||||
QY 1053 caccagaagggttttttccctacaccgcaagaacatatattagatataatctca 1112
   |||||||
DB 655 CACCAGAAAGGCTTTTTCCTCAACACAGCAAGACATATATTAGTACATGATCTCA 596
   |||||||
QY 1113 attataattatgcatatttgcattttatctcaaaataactgtggtggaatgta 1172
   |||||||
DB 595 ATTATATTTATGGCATTTATTTGCAATTTTCAAAATTAACCTGTGGGACATGTAAT 536
   |||||||
QY 1173 ctcttgagcaatctgatatcttgggaagtccttaaaagttaacaattatacaaa 1232
   |||||||
DB 535 CTCCTGAGCAATCTGATATTTTGGGAAGTCCTTTAAAAAGTTTCAATTTATCAATAAA 476
   |||||||
QY 1233 ttactgtaataagatagatctcagaacaaagaaatacagaattagagatgtcgtc 1292
   |||||||
DB 475 TTACTATGATATGATGATGATTCGAAACAAAAGAAATTCACAGATTAAGTGTGGCTGG 416
   |||||||
QY 1293 ctggtgtaagaaagcaccatgtgatgaaattcataagttgcaaaagtcacaactactgt 1352
   |||||||
DB 415 CTGCTGATGATGAGCAACATGATGATGATTTCAATAAGTTGCAAAAGTCAAAACATACTGT 356
   |||||||
QY 1353 acatgcaacagaatacaaaataataccagaatagagacctatataatgcatattata 1412
   |||||||
DB 355 ACATGCAACAGCAAGAAATCAAAATTAATCCAGAAATGAGACCTATATTAATGCAATTTAATA 296
   |||||||
QY 1413 catgaatactttgacataataagccattggaacacggaagaattagataactaataacat 1472
   |||||||
DB 295 CATGATACTTTTGACATTAATAGCCATTGGAACGGAAGATTAGATTAATAATTAACAT 236
   |||||||
QY 1473 tgactatccttggtaaatagactaaatgtagttagtactttccatggtggaat 1532
   |||||||
DB 235 TGACTATCTCTTTTAATATACAGTCACATAAATGATGTTCTTTCCATGGTGGAT 176
   |||||||
QY 1533 tttataacttttcttctgttaattttctcctgtaattttaacaataagctgtaata 1592
   |||||||
DB 175 TTTATTTACTTTTCTTTCTTTGTAATTTTCTCTCTGATATTTTAAACAATATAGCTGTATA 117
   |||||||
QY 1593 gttacaatatataaagatatgttcaaatgtgaagggcaagccaggttcagaacttt 1652
   |||||||
DB 116 GTTATCATATTTATTAAGATATTGTTCAAAATTTGAAGGGCAAAAGCCAGGTTCAGCAATTT 57
   |||||||
QY 1653 tcaaacgtatgatacttaataaataactataaataaataatataatttc 1705
   |||||||
DB 56 TCAAACTGATGTACATTTAATAAATACTATTAATTAATAATATATATTTTC 4
   |||||||

```

RESULT 8
 AAF55043
 ID AAF55043 standard; DNA; 1316 BP.
 XX AAF55043;
 AC
 XX
 DT 15-MAY-2001 (first entry)

XX Nucleotide sequence of a calcium channel transport polypeptide.
 DE
 XX
 XX Calcium channel transport polypeptide; calcium trafficking;
 KW neural disorder; HIV-induced dementia; immune system disorder;
 KW rheumatoid arthritis; muscular disorder; muscle contractile dysfunction;
 KW reproductive disorder; gastrointestinal disorder; pulmonary disorder;
 KW cardiovascular disorder; arrhythmia; renal disorder;
 KW proliferative disorder; cancer; lung carcinoma; breast cancer; ss.
 KW
 XX
 XX Homo sapiens.

OS
 XX
 XX Key Location/Qualifiers
 FH CDS 36..341
 FT /tag= a
 FT /product= "calcium channel transport polypeptide"
 XX

W0200108635-A2.

08-FEB-2001.

27-JUL-2000; 2000MO-US20392.

28-JUL-1999; 99US-0145958.

18-AUG-1999; 99US-0149446.

14-MAR-2000; 2000US-0189064.

(HUMA-) HUMAN GENOME SCI INC.

Ruben SM, NI J, Shi Y;

WPI; 2001-138604/14.

P-PSDB; AAB67466.

New invented nucleic acid useful for diagnosing, detecting, or treating

or preventing diseases associated with anomalies in calcium trafficking

across the plasma membrane -

Claim 1; Page 254; 259pp; English.

The present sequence encodes a calcium channel transport polypeptide.
 The polynucleotides, polypeptides, and antibodies are useful for
 preventing, treating, or ameliorating diseases associated with anomalies
 in calcium trafficking across the plasma membrane. They are used to
 diagnose, detect and treat or prevent diseases or conditions such as
 neural disorders (e.g. HIV-induced dementia), immune system disorders
 (e.g. Rheumatoid arthritis), muscular disorders, gastrointestinal disorders,
 dysfunction), reproductive disorders, gastrointestinal disorders, renal
 pulmonary disorders, cardiovascular disorders (e.g. arrhythmias), renal
 disorders, proliferative disorders, and/or cancerous diseases and
 conditions (e.g. lung carcinoma or breast cancer).

Sequence 1316 BP; 458 A; 227 C; 237 G; 394 T; 0 other;

Query Match 43.3%; Score 738.6; DB 22; Length 1316;
 Best Local Similarity 98.7%; Pred. No. 7.5e-159;
 Matches 755; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

```

QY 933 aaaaaaacctcaagaaagcacaataatgtctacacaaagtgcacaatttta 992
   |||||||
DB 526 aaaaaaacctcaagaaagcacaataatgtctacacaaagtgcacaatttta 585
   |||||||
QY 993 agaagcctcatgagctgtgtggtggaatcagaataaataacaggaagaaca 1052
   |||||||
DB 586 agaagcctcatgagctgtgtggtggaatcagaataaataacaggaagaaca 645
   |||||||
QY 1053 caccagaagggttttttccctacaccgcaagaacatatattagatataatctca 1112
   |||||||
DB 646 caccagaagggttttttccctacaccgcaagaacatatattagatataatctca 705
   |||||||
QY 1113 attataattatgcatatttgcattttatctcaaaataactgtggtggaatgta 1172
   |||||||

```

```

Dh 706 attataattatgcatttaatttcattttatccaataattacttggtggagatglaa 765
Qy 1173 cctctgagaactcgaattttttgggaatcctcttaaaagtacaattatcaataa 1232
Dh 766 cctctgagaactcgaattttttgggaatcctcttaaaagtacaattatcaataa 825
Qy 1233 ttaactagtagaataatgattcagaacaaagaataacagaattagatgtgctg 1292
Dh 826 ttaactagtagaataatgattcagaacaaagaataacagaattagatgtgctg 885
Qy 1293 ctggtgtagaagcaccatgtagatgattcacaagaatgcaaaagtcaaaactactg 1352
Dh 886 ctggtgtagaagcaccatgtagatgattcacaagaatgcaaaagtcaaaactactg 945
Qy 1353 acatcacaacccaataaataaataccagaataagaccattataatgattataa 1412
Dh 946 acatcacaacccaataaataaataccagaataagaccattataatgattataa 1005
Qy 1413 catgtagatttgcataataaagccattggaacggaagaattagataataaact 1472
Dh 1006 catgtagatttgcataataaagccattggaacggaagaattagataataaact 1065
Qy 1473 tgactatctcttgaataacagcactcaaatgtagttagtcttccatggtggaat 1532
Dh 1066 tgactatctcttgaataacagcactcaaatgtagttagtcttccatggtggaat 1125
Qy 1533 ttaatactcttcccttgtaatttctctctgtatatttaacaataagctgtata 1592
Dh 1126 ttaatactcttcccttgtaatttctctctgtatatttaacaataagctgtata 1184
Qy 1593 gttaacaatatataaagatttgcataatggaagcgaagccaggtccagaatt 1652
Dh 1185 gttaacaatatataaagatttgcataatggaagcgaagccaggtccagaatt 1244
Qy 1653 tcaaatgtagttagatgatttaataaataactataaataaataa 1697
Dh 1245 tcaaatgtagttagatgatttaataaataactataaataaataa 1289

```

RESULT 9

```

ABA08976
ID ABA08976 standard; cDNA; 1353 BP.
XX
AC ABA08976;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human secreted protein homologue-encoding cDNA, SEQ ID NO:752.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; growth factor;
KW inhibin; chemotaxis; chemokinesis; thrombolytic; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antitumor; ss.
XX
OS Homo sapiens.
XX
PN WO200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US03800.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.

```

XX
PA (HYSE-) HYSEQ INC.
PI Tang YT, Liu C, Drmanac RT;
XX
XX
DR MPI: 2001-457740/49.
DR P-PSDB: ABB11732.
XX
XX
PT Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PR e.g. arthritis and cancer -
XX
PS Claim 1; Page 677; 1963pp; English.

CC Sequences ABB10961-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities, stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be
CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to the treatment of viral,
CC polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a cDNA encoding a
CC novel human polypeptide of the invention.

Sequence 1353 BP; 349 A; 298 C; 315 G; 391 T; 0 other;

Query Match 40.3%; Score 687; DB 22; Length 1353;
Best Local Similarity 100.0%; Pred. No. 4.5e-147;
Matches 687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 8 gaggaagctttccactacagcgtgtatgacactggtgagtcgggccatgtagaa 67
Dh 7 gaggaagctttccactacagcgtgtatgacactggtgagtcgggccatgtagaa 66
Qy 68 atgtagcgaagataataaagcttaattgtaataataaagaataatttccct 127
Dh 67 atgtagcgaagataataaagcttaattgtaataataaagaataatttccct 126
Qy 128 tgaactatctccgtaaaagcattgtgcctcctcttggtgggacagtggttacaatcaa 187
Dh 127 tgaactatctccgtaaaagcattgtgcctcctcttggtgggacagtggttacaatcaa 186
Qy 188 tggccttgaggagccttgagtcgaagttggagagccttggagattcagatgcttcac 247

```

DB 187 tggccttgagagactcttgtagtcaagttagagccttggagattcagatgcttcaac 246
 QY 248 tggctttatctctccctctcctcgttaataatccctcctatactgtaggaactttg 307
 DB 247 tggctttatctctccctctcctcgttaataatccctcctatactgtaggaactttg 306
 QY 308 ctgcaacatctctctgtagtctgctggttcacatgctgtagcaataatactgtagctg 367
 DB 307 ctgcaacatctctctgtagtctgctggttcacatgctgtagcaataatactgtagctg 366
 QY 368 gtaataaactggaatcctcctcagtggaagatgcccctcttgagatctctacccactagact 427
 DB 367 gtaataaactggaatcctcctcagtggaagatgcccctcttgagatctctacccactagact 426
 QY 428 caaatctgagcagagagagatgctcctgctgcttgctgcatcccccagtgtagcttctacc 487
 DB 427 caaatctgagcagagagagatgctcctgctgcttgctgcatcccccagtgtagcttctacc 486
 QY 488 tgaatgagactatccacagacagaatgagagacagaacacctgtgtgagatgagctggg 547
 DB 487 tgaatgagactatccacagacagaatgagagacagaacacctgtgtgagatgagctggg 546
 QY 548 tatatgataaagtactctccctcctgagacattgtagtaagtggagctgtatgtatt 607
 DB 547 tatatgataaagtactctccctcctgagacattgtagtaagtggagctgtatgtatt 606
 QY 608 atcagtcactgaaatcagtggttcaattcctactcctgactggaatgctgtggagagga 667
 DB 607 atcagtcactgaaatcagtggttcaattcctactcctgactggaatgctgtggagagga 666
 QY 668 tcaatagtgccagatgctcctcagacaggt 694
 DB 667 tcaatagtgccagatgctcctcagacaggt 693

RESULT 10
 AA159215 standard; CDNA: 1375 BP.
 ID AA159215
 XX AA159215;
 AC
 DE 22-OCT-2001 (first entry)
 DT XX
 DE Human polynucleotide SEQ ID NO 1418.
 XX
 XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW Leukemia; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO20015312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PE 26-DEC-2000; 2000MO-US34263.
 XX
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-063450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 PA (HSE-) HXSEQ INC.
 XX
 PI Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI: 2001-442253/47.
 DR P-PSDB; AAM40059.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PS such as central nervous system injuries -
 XX
 PS Claim 1; SEQ ID NO 1418; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotide
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localized neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: immune system suppression,
 CC activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 SQ Sequence 1375 BP; 354 A; 329 C; 324 G; 368 T; 0 other;

Query Match 23.4%; Score 399; DB 22; Length 1375;
 Best Local Similarity 77.0%; Pred. No. 2,2e-81;
 Matches 486; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 64 agaaatgtagcagagatcacaatcagaacttaattgataaanaaggaatattct 123
 DB 165 agagagtgagggtcagagatcacaatcagaacttaattgataaanaaggaatattct 224
 QY 124 ccccttgaaactatcctcgttaagaagcattgctcctcctcttgagggtcagtggttcaaa 183
 DB 225 cctcttgaaactatcctcgttaagaagcattgctcctcctcttgagggtcagtggttcaaa 284
 QY 184 tcaatgaccttgaggagctcttgagaaagtggagccttgaggagattcagatgctt 243
 DB 285 tcaatgaccttgaggagctcttgagaaagtggagccttgaggagattcagatgctt 344
 QY 244 catcgttttatctctccctctcctcctcagtgtaataatccctcctatagctcagaac 303
 DB 345 cagactgtttttctcctcctcctcctcctcagtgtaataatccctcctatagctcagaac 404
 QY 304 ttgctgagcattcctcgtgtagctggttggtcagcagcttgagcaataatactgga 363
 DB 405 ttgctgagcattcctcgtgtagctggttggtcagcagcttgagcaataatactgga 464
 QY 364 tctgtaataatcagtggaatcctcagtggaatgagcctcttgagatctcattccacta 423
 DB 465 tctgtaataatcagtggaatcctcagtggaatgagcctcttgagatctcattccacta 524
 QY 424 gactcaaatcagtggaatcctcagtggaatgagcctcttgagatctcattccacta 483
 DB 525 gactcaaatcagtggaatcctcagtggaatgagcctcttgagatctcattccacta 584
 QY 484 caactgaatgagactatccacagacagaatgaggaacagaacacctgtgttgatggc 543
 DB 585 caactgaatgagactatccacagacagaatgaggaacagaacacctgtgttgatggc 644
 QY 544 tgggtatataataaagctctcctcctcagtggaatgaggaatgaggaatgaggaat 603
 DB 645 tgggtatataataaagctctcctcctcagtggaatgaggaatgaggaatgaggaat 704
 QY 604 gattatcagtaataatcagtggaatgaggaatgaggaatgaggaatgaggaatgaggaat 663
 DB 705 gattatcagtaataatcagtggaatgaggaatgaggaatgaggaatgaggaatgaggaat 764

QY 664 ggcacataggtgacatgctcagacaggt 694
 ||||| |||| | ||| ||||| |||||
 Db 765 ggcacataggtgacatgctcagacaggt 795

RESULT 11

AAD09561
 ID AAD09561 standard; cDNA; 1986 BP.

AC AAD09561;

DT 10-SEP-2001 (first entry)

DE Human transporter and ion channel-10 (TRICH-10) cDNA.

XX Human; transporter and ion channel-10; TRICH-10; cystic fibrosis; mood;
 KW gene therapy; amyotrophic lateral sclerosis; amnesia; muscular dystrophy;
 KW hypertension; angina; neurological disorder; asthma; bipolar disorder;
 KW dementia; depression; Alzheimer's disease; epilepsy; vaccine; arrhythmia;
 KW Pick's disease; Ischaemic cerebrovascular disease; AIDS; anxiety; stroke;
 KW Huntington's disease; Parkinson's disease; cerebral neoplasm; allergy;
 KW demyelinating disease; mental disorder; schizophrenia; polyomyelitis;
 KW muscle disorder; cardiomyopathy; cataract; myocarditis; Grave's disease;
 KW dermatomyositis; diabetes mellitus; immunological disorder; psoriasis;
 KW rheumatoid arthritis; Sjogren's syndrome; systemic lupus erythematosus;
 KW sickle cell anaemia; Wilson's disease; infertility; Cushing's disease;
 KW scleroderma; pulmonary artery stenosis; nontopic; Addison's disease;
 KW malabsorption syndrome; hypercholesterolaemia; cancer; ss.

OS Homo sapiens.

Key Location/Qualifiers
 CDS 248..1909

/tag=a
 /product="Human TRICH-10 protein"

MO200146258-A2.

28-JUN-2001.

22-DEC-2000; 2000MO-US35095.

23-DEC-1999; 98US-0172000.

PR 14-JAN-2000; 2000US-0176083.

PR 21-JAN-2000; 2000US-0177332.

PR 28-JAN-2000; 2000US-0178572.

PR 02-FEB-2000; 2000US-0179758.

PR 10-FEB-2000; 2000US-0181625.

PA (INCY-) INCYTE GENOMICS INC.

PI Baughn MR, Burford N, Au-Young J, Lu DM, Yang J, Reddy R, Lal P;

PI Hillman JL, Azimzai Y, Yue H, Nguyen DB, Yao MG, Gandhi AR;

PI Tang YT, Khan FA;

XX WPI; 2001-418042/44.

DR P-PSDB; AAE04897.

PT Novel human transporter and ion channel proteins useful for treating

PT and preventing transport, neurological, muscle and immunological

XX disorders

XX Claim 5; Page 147; 160pp; English.

XX The present sequence is transporter and ion channel-10 (TRICH-10) cDNA.

XX TRICH is used as vaccine. TRICH is useful for treating a disease or

XX condition associated with decreased expression of functional TRICH,

XX such as transport disorder including amyotrophic lateral sclerosis,

XX cystic fibrosis, Becker's muscular dystrophy, Charcot-Marie Tooth

XX disease, Duchenne muscular dystrophy, angina and hypertension,

XX neurological disorders including Alzheimer's disease, amnesia, bipolar

XX disorder, dementia, depression, epilepsy, ischaemic cerebrovascular

CC disease, stroke, cerebral neoplasms, Pick's disease, Huntington's
 CC disease and Parkinson's disease, schizophrenia, demyelinating diseases, mental disorders
 CC including mood, anxiety, depression, asthma, bipolar disorder,
 CC muscle disorder including cardiomyopathy, myocarditis, polyomyelitis,
 CC dermatomyositis, arrhythmias and asthma and immunological disorders
 CC including AIDS, adult respiratory distress syndrome (ARDS), allergies,
 CC anaemia, diabetes mellitus, rheumatoid arthritis, scleroderma, Sjogren's
 CC syndrome, systemic lupus erythematosus and other diseases including
 CC sickle cell anaemia, Wilson's disease, cataracts, infertility, pulmonary
 CC artery stenosis, Grave's disease, Cushing's disease, Addison's disease,
 CC glucose-galactose malabsorption syndrome, hypercholesterolaemia, cancers
 CC psoriasis and viral, bacterial, fungal, helminthic and protozoal
 CC infections. TRICH DNA is useful in gene therapy and in diagnostic
 CC purposes.
 XX

SO Sequence 1986 BP; 505 A; 496 C; 457 G; 528 T; 0 other;

Query Match 23.4%; Score 399; DB 22; Length 1986;

Best Local Similarity 77.0%; Pred. No. 2,36-81;

Matches 486; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 64 agaatgtgagcagagataacacagcttaattgaattaaaggaattattct 123

Db 125 agagagtggtggtcagatcaacacatttagtgacttggagaaacacatttcc 184

QY 124 ccttgaactatctcgttaagcatttgccctcttgggggtacagtggtacaa 183

Db 185 ccttgaactatctcgttaagcatttgccctcttgggggtacagtggtacaa 244

QY 184 tcaatgaccttgagagccttgatcaagttgagagccttgagatccatgctt 243

Db 245 tcaatgaccttgagagccttgatcaagttgagagccttgagatccatgctt 304

QY 244 catcgttttattctcctcctcctcctcctcctcctcctcctcctcctcct 303

Db 305 cagactgttttctcctcctcctcctcctcctcctcctcctcctcctcctcct 364

QY 304 ttctgtagcaccatctcctcctcctcctcctcctcctcctcctcctcctcct 364

Db 365 ttctgtagcaccatctcctcctcctcctcctcctcctcctcctcctcctcct 424

QY 364 tctgtagcaccatctcctcctcctcctcctcctcctcctcctcctcctcct 423

Db 425 tctgtagcaccatctcctcctcctcctcctcctcctcctcctcctcctcct 484

QY 424 gactcaaatctgagcagagagtgctgctgttgcattccatccagtgagctt 483

Db 485 gactcaaatctgagcagagagtgctgctgttgcattccatccagtgagctt 544

QY 484 cactgaatgagcactccacacagcagagtgagcagacacacacacacacac 543

Db 545 cactgaatgagcactccacacagcagagtgagcagacacacacacacacac 604

QY 544 tgggtatgatcaaaagctactccttcagacattgagcagagtgagcagctt 603

Db 605 tgggtatgatcaaaagctactccttcagacattgagcagagtgagcagctt 664

QY 604 gattatcagctaatcaatcagtggttaattcctcctcctcctcctcctcct 663

Db 665 gattatcagctaatcaatcagtggttaattcctcctcctcctcctcctcct 724

QY 664 ggcacataggtgacatgctcagacaggt 694

Db 725 ggcacataggtgacatgctcagacaggt 795

XX

XX

XX

XX

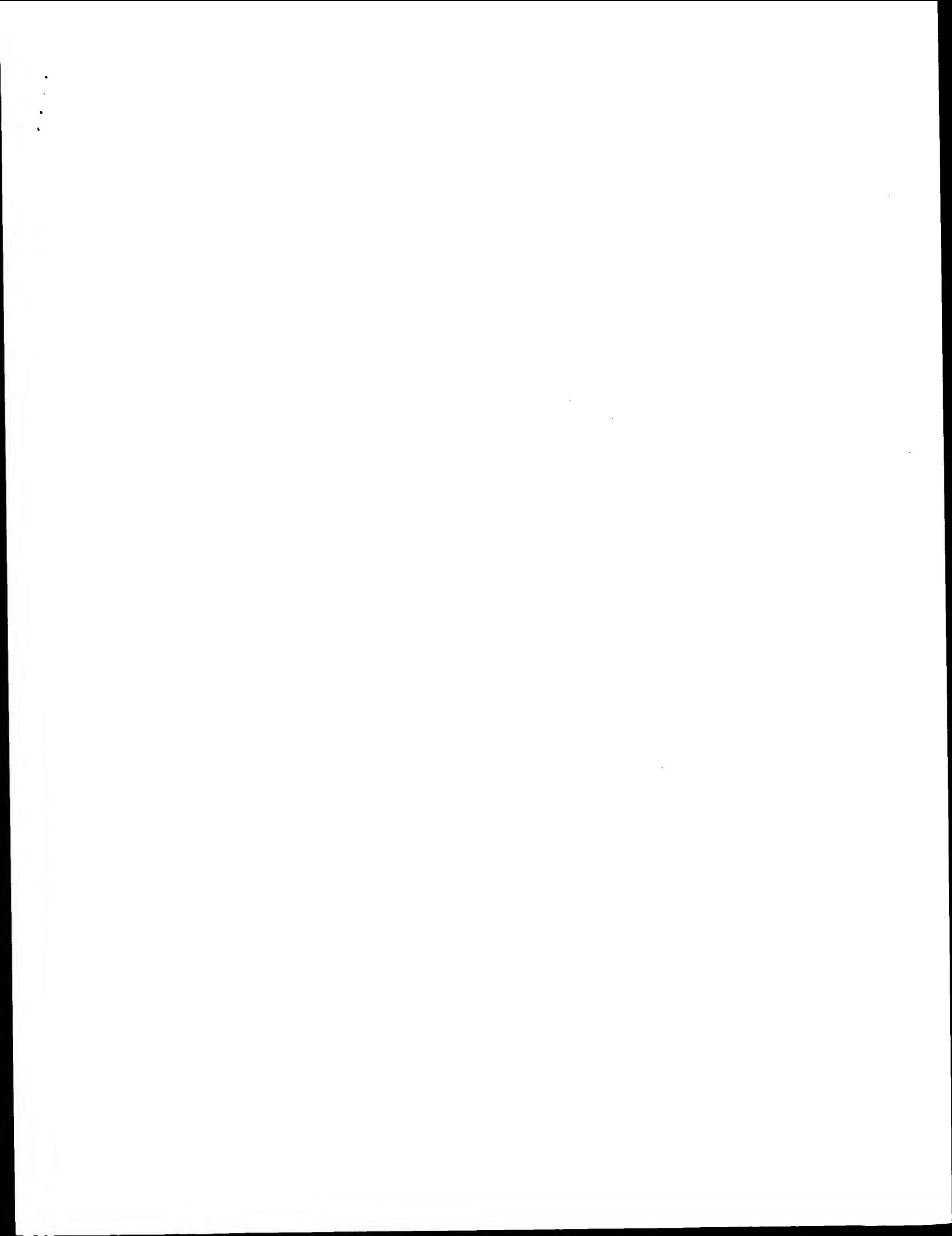
XX

XX

XX

XX

XX



	RESULT	1		DNA	linear	PAT 06-FEB-2001
	Locus	AX074150	2684 bp			
	DEFINITION	Sequence 6 from Patent WO0104283.				
	ACCESSION	AX074150				
	VERSION	AX074150.1	GI:12710362			
	KEYWORDS					
	SOURCE	human,				
	ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 2684)				
	REFERENCE	Sun,W. Organic anion transporter genes and proteins Patent: WO 0104283 A 6 18-JAN-2001;				
	AUTHORS	JOURNAL	Metabasis Therapeutics Inc. (US) Location/Qualifiers			
	FEATURES	source				
		..2684 '/organism='Homo sapiens' /db_xref='taxon:9606' 810 t 533 c 541 g				
	BASE COUNT	800 a				
	ORIGIN					

Query Match 47.0% Score 801.8; DB 6; Length 2684;
 Best Local Similarity 66.2% Pred. No. 1.3e-146;
 Matches 1693; Conservative 0; Mismatches 12; Indels 854; Gaps 3;

1 aagaactgaggaagctcttcacacagctgtatgcaactggtgagatccgggcccattg 60
 14 AAGAAGTAGAGAAAGCTTTCCACACGCGTGTATGCTACCTGGAGTGGAGCCCATG 73
 61 atgagaatgatgcsagagatcaatacaagcttaattgtaataataaaggaaatatt 120
 74 ATGAGAAATGATGCGAGAGATCAATACAACTTAATTGATTAATAAAGAAATATT 133
 121 tctcccttgaactatctccgtaagccattgscctctctgggggtcaggttca 180
 134 TCTCCCTTGAACCTTATCTCCGTAAGCCATTGTGCTCTCTTGGGGGTCACTGTTCA 193
 181 caatcaatggccttgaaggagctctgagcaagtctgagagccttggagattcaagt 240
 194 CAATCAATGGCCTTGAGAGCTCTTGAGTCAAGTTGGAGGCTTGGAGATTTCAGATG 253
 241 ctcaactggttttattctctccctctctcattgtatataatccatatactagag 300
 254 CTCAACTGGTTTATTATCTCTCCCTCTCATGTTTAAATCCCTCATATAGCTAGAG 313
 301 aacttctgacagcaatctcctgctcactgctgctgaggtccacatgctgagaa 360
 314 AACTTCTGACAGCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 373
 361 ggaactggttaagaaatcgaatccatgagaaatgagcctcttggagaaatccat 420
 374 GGAACTGTTAAGAACTGAAATCCTGAGTGAAGAGATGCCCTGAGATCTCTATCCA 433
 421 ctgaactcaatctgagggccagagagtgctgctgctgctgctgctgctgctgct 480
 434 CTGAACTCAATCTGAGGCGCAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 493
 481 ctgaactcaatctgagggccagagagtgctgctgctgctgctgctgctgctgct 540
 494 CTGAACTCAATCTGAGGCGCAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 553
 541 ggcctggtatataagcaagctctcctcctgacacattgtaagaaatggagcct 600
 554 GCGCTGGTATATGATCAAAAGCTACTCTCTGACCACTGCTGACTAAGTGGAGCTG 613
 601 tctgattatcagctacagaaatcaagtggttcaatctcctgctgctgctgctgct 660
 614 TGTGATTATCAGTCACTGAAATCAGTGTTCATTTCTTACTGTGAGTGAATGCTG 673
 661 ggaagcctcaatgagtgagcctgctgctgctgctgctgctgctgctgctgctg 690
 674 GGAAGCCTCAATGAGTGGCATGTCTGACAGAGTGGGCAAGATTTATTTCTCATG 733
 691 ----- 690
 734 TGTGCTCCAGCTTGCCATTACTGACACCTGCGCTGCTTCCGCTCCACCTTCCCTG 793
 691 ----- 690
 794 TACTGTACTACGCTTCTTGGAGGTTTTTCTTCATGATCATATATCAATAATATT 853
 691 ----- 690
 854 TTGCCCATTACTAGTAGATAGAGCCCACTTAAGCCCTGTAGTAAATATTGATCT 913
 691 ----- 690
 914 GGTGCCCTTAGTATTGGACAGATATCTGGAGGCTTGGCTTATGTTCCGAGACTGG 973
 691 ----- 690
 974 CAACCTGACGCTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1033

QY 697 ctggtgaatctgtcctcgtggtgtaatacaaccaataaactagatggagcttaagga 756
 DB 1034 CTGCTGGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 756
 QY 757 cttaaaaaagttgcaagcaaatgaaatgaagaaatgctgaagaaacccctgaagga 816
 DB 1094 CTTAAGAAAGTTGCAAGCAGCAAAATGAAATGAAGTCTGGAAGAAACCTGAACTAGAG 1153
 QY 817 gttgtaagatccaccatgcaagagagagagctgagatgcaagcaagcaagaaactctgtgt 876
 DB 1154 GTTGAAGATCCACAGATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1213
 QY 877 gacttctccgaaccccgatgctgtaagagatctgtatcttggag----- 932
 DB 1214 GACTTGTCCGCAACCCCGATGAGTAAAGAGATCTGATCTGCTGCTGCTGCTGCTG 1273
 QY 933 ----- 932
 DB 1274 GCAAAACAAATACCTTTTATGTTACATGTCATCTTACAGATGAGGAGCAACATT 1333
 QY 933 ----- 932
 DB 1334 TTCTGTTGCAAGTACTTATGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1393
 QY 933 ----- 932
 DB 1394 ACATTAATCATATGAGGCGCTGCAATAGCCAGATTTGTCATGTTCTGCTGCTGCTG 1453
 QY 933 ----- 932
 DB 1454 TCCATTTTGGCAACACGTTTGGCCCAAGAAATGACAGACCCTGCTGCTGCTGCTG 1513
 QY 933 ----- 932
 DB 1514 TGTCTGGAATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1573
 QY 933 ----- 932
 DB 1574 CTGATCCCACTGTTCTGAGGCAAGAGCTTCAAGATATTAAGGCTAGTAGATT 1633
 QY 933 ----- 932
 DB 1634 GGAGCAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1693
 QY 933 ----- 932
 DB 1694 ATCATTTATGAAATCTTCCCATCATGTTGGCTTATGCTTCTCTCTACAGAAAC 1753
 QY 933 ----- 932
 DB 1754 AAGAATCTGCTTGGCTGACACCATCAAGATGTGAAATCAAAAAAATCTCAAG 1813
 QY 948 gaaaagcaataaagtgttcttaacaaagtgaacaaatcttaagaaaccttcaag 1007
 DB 1814 GAAAAGCATTAATAATGATGCTGACCAAAAGTACCAAAATTTAAGAACTCTCAAG 1873
 QY 1008 ctgattgtgggaaatctcagaaaaaaatacaagagagagagagagagagagagag 1067
 DB 1874 CTGATGTGTGGGAAATTCAGAAAAAATAACAGGAAAAAGAACACCAAGAGGCTTT 1933
 QY 1068 tttccctacaacagcaagacatatatagataatgaatcgaatctcaatlaaataagga 1127
 DB 1934 TTTCCCTACAAACAGCAAGAAATATATATGATGATGATGATGATGATGATGATGAT 1187
 QY 1128 ttaattgcaattatlaaataaactgctgagagagagagagagagagagagagagag 1247
 DB 1994 TTAATTTGCAATTTATTAATAATTAATTTGAGGAGATGATGATGATGATGATGAT 2053
 QY 1188 atattttggagagctcttaaaaaagttacaatattcaataaacttaagtaagaa 1247
 DB 2054 ATATTTTTGGAGGCTCTTAAAGTTTACAAATTTATCAATTAATTAATTAATTAAT 2113
 QY 1248 atgattcagaacaaagaatacaagaaatgaagatgctgctgctgctgctgctgctgctg 1306

```

ALL repeats were identified using RepeatMasker
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RW/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www.seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information

----- Center project name: L12735
Center clone name: 3110_P_2
Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 157859 bases at least Q40
Consensus quality: 151528 bases at least Q30
Consensus quality: 162934 bases at least Q20
Insert size: 172000; agarose-fp
Insert size: 163400; sum-of-contigs
Quality coverage: 5.0 in Q20 bases; sum-of-contigs
Quality coverage: 5.2 in Q20 bases; sum-of-contigs

-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces is
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 1069: contig of 1069 bp in length
*
* 1070 1169: gap of 100 bp
*
* 1170 2379: contig of 1210 bp in length
*
* 2380 2479: gap of 100 bp
*
* 2480 3441: contig of 961 bp in length
*
* 3441 3540: gap of 100 bp
*
* 3541 4880: contig of 1340 bp in length
*
* 4881 4980: gap of 100 bp
*
* 4981 6766: contig of 1788 bp in length
*
* 6769 6868: gap of 100 bp
*
* 6869 8318: contig of 1450 bp in length
*
* 8319 8418: gap of 100 bp
*
* 8419 10046: contig of 1628 bp in length
*
* 10047 10146: gap of 100 bp
*
* 10147 12455: contig of 2309 bp in length
*
* 12456 12553: gap of 100 bp
*
* 12556 15164: contig of 2609 bp in length
*
* 15165 15264: gap of 100 bp
*
* 15265 16648: contig of 1384 bp in length
*
* 16649 16748: gap of 100 bp
*
* 16749 29429: contig of 12681 bp in length
*
* 29430 29529: gap of 100 bp
*
* 29530 41226: contig of 11697 bp in length
*
* 41227 41326: gap of 100 bp
*
* 41327 57370: contig of 16044 bp in length
*
* 57371 57470: gap of 100 bp
*
* 57471 102968: contig of 45498 bp in length
*
* 102969 103068: gap of 100 bp
*
* 103069 164809: contig of 61741 bp in length.
*
Location/Qualifiers
1.164809
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="CTD-3110P2"
/clone_lib="CITD Hunan BAC"
1.1069
/misc_feature
/note="assembly_fragment"
1170.2379
/misc_feature
/note="assembly_fragment"

```

```

misc_feature      2480..3440
                    /note="assembly-fragment"
misc_feature      3541..4880
                    /note="assembly-fragment"
misc_feature      4981..6768
                    /note="assembly-fragment"
misc_feature      6869..8318
                    /note="assembly-fragment"
misc_feature      8419..10046
                    /note="assembly-fragment"
misc_feature      10147..12453
                    /note="assembly-fragment"
misc_feature      12556..15164
                    /note="assembly-fragment"
misc_feature      15265..16648
                    /note="assembly-fragment"
misc_feature      16749..29429
                    /note="assembly-fragment"
misc_feature      29530..41226
                    /note="assembly-fragment"
misc_feature      41327..57370
                    /note="assembly-fragment"
misc_feature      57471..102968
                    /note="assembly-fragment"
misc_feature      103069..164809
                    /note="assembly-fragment"
BASE COUNT      50889 a 32076 c 31689 g 48754 t 1401 others
ORIGIN

```

```

Query Match      45.4%; Score 774.8; DB 2; Length 164809;
Best Local Similarity 99.1%; Pred. No. 1.5e-141;
Matches 779; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 920 tggatatttgaagaaaaaaatcccaaggaagcctaaatgaatgctacacaaag 979
    ||| |||
Db 26917 TTTGTTTTCAGAAAAAAATCTCAGAGAAAGCAATTAAGATTCCTACACAAAG 26858
    ||| |||

Qy 980 tgacaaatttaagagagccttcagagctgagtggtggtgaatcagaagaaaaata 1039
    ||| |||
Db 26857 TGACCAATTTTAAAGAGCCCTCATGAGCTGATGAGGGAATTCAGAAAAAAATA 26798
    ||| |||

Qy 1040 caggaagaaagacacacagagaggttttccctcaacccgcaagacatatata 1099
    ||| |||
Db 26797 CAGGAAAGAACACACAGAGAGGTTTTTCCCTCAACACGCAAGACATATATTAGA 26738
    ||| |||

Qy 1100 tacatgaatcgaattatataatgagcaatattgcaatttatccaataactgt 1159
    ||| |||
Db 26737 TACATGAATCTCATTTATATATATGCAATTTTGCATTTTATTTCAAAATTAATTGT 26678
    ||| |||

Qy 1160 gggagacatgaatcctcttgagcaatcgtatatttgggaagccttaaaagtaca 1219
    ||| |||
Db 26677 GGGGACATGATATCTCTTGAGCAATCTGATATTTTGGGAAGTCTTTAAAAAGTTACAA 26618
    ||| |||

Qy 1220 attatcaataatattactaagatagaatgaatgaagaacagaatcacagaatt 1279
    ||| |||
Db 26617 ATTATCAATAATATTACTACTAGTAGATTAAGATGATTCAGAAACAAAAGAAATCACAGAAAT 26558
    ||| |||

Qy 1280 aggatgtgagctgctgglytgaagacacatgagatgaatgaatgaatgaatga 1339
    ||| |||
Db 26557 AGGATGTGCTGCTGGCTGTATGAAGACCATGATGAATTCATTAAGTTGCAAAAAGTC 26498
    ||| |||

Qy 1340 aaacaatattgcatgagcaacagaaatcaaaaataatcagaataatgagactata 1399
    ||| |||
Db 26497 AAAACATATTGTCATGACAGCAAGAAATCAAAATTAATTCAGAAATGAGACCTTATA 26438
    ||| |||

Qy 1400 aatgcatatgaatgaatccttgacataaagccattggaagaaaggaagattaga 1459
    ||| |||
Db 26437 AATGCAATTAATGATGATGATTTTACATTAATTAAGCCATTTGGAAGAAACGGAAGATTAGA 26378
    ||| |||

Qy 1460 taataataactgactctctcttgtaataacagtcactaataatgagtgttgtaactt 1519
    ||| |||
Db 26377 TACTAATAATCACTTGACTATCTCTTTGTAATAAGAGTACTAATAATGATGTTTACTTT 26318
    ||| |||

```

```

Qy 1520 tccatggtggaatttaattactttctcttgtaattttctctgatatatttaaca 1579
    ||| |||
Db 26317 TCCATGCTGGAATTTTAAATTAATCTTTCTTGTAAATTTTCTCTGTAATTTTAAACA 26258
    ||| |||

Qy 1580 aatactgcatagttacacatatataaagatattgctcaattgaaaggcaaggcca 1639
    ||| |||
Db 26257 AATAGCTGTATAGTTTACATATTATAAGATATTGTTCAATTAAGGCAAGGCCA 26198
    ||| |||

Qy 1640 ggtcagcaattttcaactgcatgacattcaataaataactaataaataaata 1699
    ||| |||
Db 26197 GGTTCAGCAATTTTCAAACTGTATGATTAATAAATACTAATTAATTAATAATTA 26138
    ||| |||

Qy 1700 tatttc 1705
    ||| |||
Db 26137 TATTTC 26132
    ||| |||

```

```

RESULT 3
AP001880 169776 bp DNA linear PRI 26-APR-2001
LOCUS Homo sapiens genomic DNA, chromosome 11q, clone:CTD-3110P2,
DEFINITION complete sequences.
ACCESSION AP001880
VERSION AP001880.4 GI:13810518
KEYWORDS HTG.
SOURCE Homo sapiens DNA, clone:CTD-3110P2.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (sites)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens genomic DNA
Published Only in Database (2000) In press
2 (bases 1 to 169776)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (24-APR-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho,Tsukumi-Ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsr.riken.go.jp, URL:http://hnp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
On Apr 26, 2001 this sequence version replaced gi:10716820.
FEATURES
source
location/Qualifiers
1..169776
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="CTD-3110P2"

```

```

BASE COUNT      48281 a 32818 c 33507 g 55170 t
ORIGIN

```

```

Query Match      45.4%; Score 774.8; DB 9; Length 169776;
Best Local Similarity 99.1%; Pred. No. 1.4e-141;
Matches 779; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 920 tggatatttgaagaaaaaaatcccaaggaagcctaaatgaatgctacacaaag 979
    ||| |||
Db 82466 TTTGTTTTCAGAAAAAAATCTCAGAGAAAGCAATTAAGATTCCTACACAAAG 82525
    ||| |||

Qy 980 tgacaaatttaagagagccttcagagctgagtggtggtgaatcagaagaaaaata 1039
    ||| |||
Db 82526 TGACCAATTTTAAAGAGCCCTTCATGAGCTGATGAGGGAATTCAGAAAAAAATA 82585
    ||| |||

Qy 1040 caggaagaaagacacacagagaggttttccctcaaacagcaagacatatata 1099
    ||| |||
Db 82586 CAGGAAAGAACACACAGAGAGGTTTTTCCCTCAACACGCAAGACATATATTAGA 82645
    ||| |||

Qy 1100 tacatgaatcgaattatataatgagcaatattgcaatttatccaataactgt 1159
    ||| |||

```


Center code: RIKEN
Web site: <http://hgp.gsc.riken.go.jp/hattori@gs.riken.go.jp>

Center	Clone	Summary	Statistics
-----	-----	-----	-----
Sequencing vector:		PCR products:	100% of reads
			100% of reads

```

Sequence: 165530 bases at least Q20
Chemistry: Dye-terminator ET-amc-shadown
Assembly program: Phrap; version 0.990329
Consensus quality: 162010 bases at least Q40
Consensus quality: 164639 bases at least Q30
Consensus quality: 165530 bases at least Q20
Consensus quality: sum-of-confids
Insert size: 165046; sum-of-confids

```

Quality coverage. 0.98
NOTE: This is a 'working draft' sequence. It currently consists of 9 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

	39293 contig of	39293 bp in length
1	39293 contig of	39293 bp in length
		39293 bp in length

```

1. .166848
/organism="Homo sapiens"

```

```

/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-151E18"

```

```
misc_feature      1. 39293
                  /note="assembly_fragment"
00304             73653
```

```
misc_feature      39354 . 1000  
/note="assembly_fragments"  
73754 . 99297
```

```
misc_feature
/note="assembly-fragment"
99398. .123005
```

```
misc_feature    /note="assembly_fragment"  
123106.  .140640
```

----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)


```

misc_feature /note="assembly-fragment"
140741..152854
misc_feature /note="assembly-fragment"
152855..161965
misc_feature /note="assembly-fragment clone_end:T7 vector_side:left"
162066..165410
misc_feature /note="assembly-fragment clone_end:SP6 vector_side:left"
165511..166848
BASE COUNT 55669 a 32391 c 31923 g 46064 t 801 others
ORIGIN

```

```

Query Match 4.7% Score 762.8; DB 2; Length 166848;
Best Local Similarity 99.0%; Pred. No. 3.2e-139;
Matches 778; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 920 tggatttttgaagaaatctcaagaaagcattaaatgatgtgtacacaaag 979
DB 36630 TGTGTTTTCAGAAAAAAATCTCAGAGAAAGCATTAATAATGATGCTACACAAAG 36571
QY 980 tggcaaatcttgaagaccccttcattgctgattggtggtggaatcagaagaaataa 1039
DB 36570 TGACCAAAATTTTAAAGAGCCCTTACAGCTGATGCTGATGCTGGAATTCAGAAAAATTA 36511
QY 1040 caagaaagaacacacacagagaggttttttccctacacacagcagacacataataga 1099
DB 36510 CAGGAAAAAGAACACACAGAGAGGTTTTCCTACACACAGACAGACATTAATATAGA 36451
QY 1100 taactgaatcctaataataatattgacattatgacattatcttcaataataactgt 1159
DB 36450 TACATGAAATCTCAATATATATATGACATTAATTTGCAATTTATTTCAAAATTAATG 36391
QY 1160 ggggaacatgtaactcttgaagcaatctgatacttttgggaagtccttaagaattaca 1219
DB 36390 GGGGACATGTAATCTCTGACCAATCTGATATTTTGGGAAGTCCTTAAAGATTTACA 36331
QY 1220 atttcaataataataataataataataataataataataataataataataataata 1279
DB 36330 ATTTATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 36271
QY 1280 aggtatggtcgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1339
DB 36270 AGGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1339
QY 1340 aaaaactactgacatgacacacacacacacacacacacacacacacacacacacac 1399
DB 36210 AAAACATATCTGATGACACACACACACACACACACACACACACACACACACACAC 1399
QY 1400 aatgcaattatgacatgacatgacatgacatgacatgacatgacatgacatgacat 1459
DB 36150 AATGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1459
QY 1460 tacttaataataataataataataataataataataataataataataataataata 1519
DB 36090 TACTTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 1519
QY 1520 tcaatggtggaatttaataataataataataataataataataataataataataata 1579
DB 36030 TCCATGCTGGAATTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1579
QY 1580 aatagctggtatagtttcaataataataataataataataataataataataata 1639
DB 35971 AATAGCTGATATAGTTTCAATATATTAATAATTAATTAATTAATTAATTAATTAAT 1639
QY 1640 ggttcaagaatttcaaacctatgtacatttaataataataataataataataata 1699
DB 35911 GGTTCAGCAATTTCAAACTATGATACATTTAATAATAATAATAATAATAATAATAATA 1699
QY 1700 tatlct 1705
DB 35851 TATTTT 35846

```

```

RESULT 5
AC073256/c
LOCUS
DEFINITION
AC073256 Homo sapiens chromosome 11 clone RP11-204K20, WORKING DRAFT
ACCESSION
AC073256
VERSION
AC073256.2 GI:8569847
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 170682)
Waterson, R.H.
The sequence of Homo sapiens clone
2 (bases 1 to 170682)
Waterson, R.H.
Direct Submission
Submitted (12-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108 USA
On Jun 16, 2000 this sequence version replaced gi:8468975.

```

COMMENT

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0204K20
----- Summary Statistics -----
Sequencing vector: M13; 100%
Chemistry: Dye primer ET; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 157404 bases at least Q40
Consensus quality: 160944 bases at least Q30
Insert size: 161000; agarose-fp
Insert size: 168782; sum-of-ctligs
Quality coverage: 4.62 in Q20 bases; agarose-fp
Quality coverage: 4.50 in Q20 bases; sum-of-ctligs
----- NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
-----
1 2201: contig of 2200 bp in length
* 2300: gap of unknown length
* 3481: contig of 1181 bp in length
* 3582: gap of unknown length
* 4832: contig of 1251 bp in length
* 4932: gap of unknown length
* 6631: contig of 1698 bp in length
* 6730: gap of unknown length
* 11545: contig of 4815 bp in length
* 11645: gap of unknown length
* 11646: contig of 4171 bp in length
* 15817: gap of unknown length
* 15917: contig of 3982 bp in length
* 19898: gap of unknown length
* 19999: contig of 4952 bp in length
* 24950: gap of unknown length
* 24951: contig of 5299 bp in length
* 30349: gap of unknown length
* 30350: contig of 6012 bp in length
* 30450: gap of unknown length
* 36462: gap of unknown length

```


TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
Homo sapiens 116,273 genomic DNA of 11q13
2 (bases 1 to 116273)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (12-OCT-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
Japan (E-mail: hattori@gs.c.riken.go.jp,
URL: http://hpg.gs.c.riken.go.jp/
Fax: 81-42-778-9924)
On May 31, 2000 this sequence version replaced gi:6997486.

Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hpg.gs.c.riken.go.jp/
Contact: hattori@gs.c.riken.go.jp
Project Information
Center project name: Xpac-662A14
Center clone name: Xpac-662A14

Sequencing Vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 103806 bases at least Q40
Consensus quality: 11869 bases at least Q30
Insert size: 113473; sum-of-coverage
Quality coverage: 4.03x in Q20 bases; sum-of-coverage

NOTE: This is a 'working draft' sequence. It currently consists of
29 contigs. The true order of the pieces is not known and the
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs 'N', but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved

```

1 38138 contig of 38138 bp in length
38239 44698 contig of 6460 bp in length
44799 49514 contig of 4716 bp in length
49615 55534 contig of 5920 bp in length
55635 60136 contig of 4502 bp in length
60237 65623 contig of 5387 bp in length
65724 70102 contig of 4379 bp in length
70203 71387 contig of 1185 bp in length
71488 73839 contig of 2352 bp in length
73940 77366 contig of 3427 bp in length
77467 80065 contig of 2599 bp in length
80166 83925 contig of 3660 bp in length
83926 86777 contig of 2852 bp in length
86878 89783 contig of 2906 bp in length
89884 92751 contig of 2868 bp in length
92852 95068 contig of 2217 bp in length
95169 96739 contig of 1571 bp in length
96840 97912 contig of 1732 bp in length
97913 99744 contig of 100 bp in length
99845 101153 contig of 1309 bp in length
101154 102523 contig of 1615 bp in length
102524 104921 contig of 1953 bp in length
104922 105021 contig of 1479 bp in length
105022 106500 contig of 1479 bp in length
106501 108529 contig of 1929 bp in length
108530 108629 contig of 100 bp in length
108630 110841 contig of 2212 bp in length
110842 110941 contig of 100 bp in length
110942 112367 contig of 1426 bp in length
112368 112467 contig of 100 bp in length
112468 113665 contig of 1198 bp in length
113666 113765 contig of 100 bp in length
113766 114974 contig of 1209 bp in length
114975 115074 contig of 100 bp in length
115075 116273 contig of 1199 bp in length
Sequence updated (26-May-2000).

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 29 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as

FEATURES

```

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
1 38138: contig of 38138 bp in length
38139 38238: gap of 100 bp
38239 44698: contig of 6460 bp in length
44699 44798: gap of 100 bp
44799 49514: contig of 4716 bp in length
49515 49614: gap of 100 bp
49615 55534: contig of 5920 bp in length
55535 55634: gap of 100 bp
55635 60136: contig of 4502 bp in length
60137 60236: gap of 100 bp
60237 65623: contig of 5387 bp in length
65624 65723: gap of 100 bp
65724 70102: contig of 4379 bp in length
70103 70202: gap of 100 bp
70203 71387: contig of 1185 bp in length
71388 71487: gap of 100 bp
71488 73839: contig of 2352 bp in length
73840 73939: gap of 100 bp
73940 77366: contig of 3427 bp in length
77367 77466: gap of 100 bp
77467 80065: contig of 2599 bp in length
80066 80165: gap of 100 bp
80166 83925: contig of 3660 bp in length
83926 86777: gap of 100 bp
86778 86777: contig of 2852 bp in length
86778 86777: gap of 100 bp
86778 89783: contig of 2906 bp in length
89784 92751: contig of 2868 bp in length
92752 92851: gap of 100 bp
92852 95068: contig of 2217 bp in length
95069 95168: gap of 100 bp
95169 96739: contig of 1571 bp in length
96740 96839: gap of 100 bp
96840 97912: contig of 1732 bp in length
97913 98012: gap of 100 bp
98013 99744: contig of 1309 bp in length
99745 99844: gap of 100 bp
99845 101153: contig of 1615 bp in length
101154 102523: gap of 100 bp
102524 104921: contig of 1953 bp in length
104922 105021: gap of 100 bp
105022 106500: contig of 1479 bp in length
106501 108529: gap of 100 bp
108530 108629: contig of 1929 bp in length
108630 110841: gap of 100 bp
110842 110941: contig of 2212 bp in length
110942 112367: gap of 100 bp
112368 112467: contig of 1426 bp in length
112468 113665: gap of 100 bp
113666 113765: contig of 1198 bp in length
113766 114974: gap of 100 bp
114975 115074: contig of 1209 bp in length
115075 116273: gap of 100 bp
Location/Qualifiers
1.116273
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
/clone="Xpac-662A14"
1.38138
/note="assembly-fragment"
38239.44698
/note="assembly-fragment"
44799.49514

```

misc_feature
misc_feature
misc_feature
misc_feature

Query Match	44.6%	Score 759.6	DB 2	Length 116213
Best Local Similarity	98.7%	Pred. NO. 1.4e-138		
Matches 776	Conservative 0	Mismatches 9	Indels 1	Gaps 1
QY 920	tgttatttttgagaaaaaaatctcaagaaagacataaataatgatttgtacaaaag	979		
Db 57920	ttcttttttccagaaaaaaatctctaagaaagacataaataatgatttgtacaaaag	57861		
QY 980	tgcacaattttaagaagcctcatgagctgatttgytgggaaatccaagaaaaata	1039		
Db 57860	tgacccaatttttaagagcctcatgagccttcaagagctgatttgytgggaaatccaagaaaaata	57801		
QY 1040	caggaaagaacacacaccagagggttttttccctacaacagcaagaagacatatatta	1099		
Db 57800	caggaagaagacacacaccagagggttttttccctacaacagcaagaagacatatatta	57741		
QY 1100	tacatgaattccaattataattatgagcataatttgatatttattccaataacttgt	1159		
Db 57740	tacatgaattccaattataattatgagcataatttgatatttattccaataacttgt	57681		
QY 1160	ggggacactgtaactctcttgagcaattgatatitttgggaagtcctttaaaagtacaa	1219		
Db 57680	ggggacactgtaactctcttgagcaattgatatitttgggaagtcctttaaaagtacaa	57621		
QY 1220	atttacaataaattactagtagataagatgattcagaaacaaaagaaatccaagaatt	1279		
Db 57620	atttacaataaattactagtagataagatgattcagaaacaaaagaaatccaagaatt	57561		
QY 1280	aggaatgcctgcctggttgtatgaagcacacvtgatgaattcataaagtctgcaaatgc	1339		

```

RESULT 7
AP003420      160945 bp      DNA      linear      PRI 28-AUG-2001
LOCUS
DEFINITION    Homo sapiens genomic DNA, chromosome 11q, clone:RP11-614K12,
               complete sequence.
ACCESSION     AP003420
VERSION       AP003420.2      GI:15320509
KEYWORDS      HTG.
SOURCE        Homo sapiens DNA, clone:RP11-614K12.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE     1 (bases 1 to 160945)
AUTHORS       Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
               Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
               Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
               Homo sapiens genomic DNA
               Published Only in database (2001) In press
TITLE         2 (bases 1 to 160945)
JOURNAL       Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
               Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
REFERENCE     Direct Submission
AUTHORS       Submitted (16-MAR-2001) Masahira Hattori, The Institute of Physical
               and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
               1-7-22 Suenhiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
               (E-mail:hattori@gsc.riken.go.jp, URL:http://hgc.gsc.riken.go.jp/)
               Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT       On Aug 27, 2001 this sequence version replaced gi:13383333.
FEATURES
SOURCE        location/Qualifiers
               1..160945
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /chromosome="11"
               /map="11q"
               /clone="RP11-614K12"
BASE COUNT    46743 a 31738 c 32047 g 50417 t
ORIGIN

```

Query Match	34.5%	Score 588;	DB 9;	Length 160945;
Best Local Similarity	100.0%;	Pred. No. 3.8e-105;		
Matches 588;	Conservative	0;	Mismatches 0;	Indels 0;
			Gaps	0

FEATURES	SOURCE
*	1
*	38138: contig of 38138 bp in length
*	38139 38238: gap of 100 bp
*	38239 44698: contig of 6460 bp in length
*	44699 44798: gap of 100 bp
*	44799 49514: contig of 4716 bp in length
*	49515 49614: gap of 100 bp
*	49615 55534: contig of 5920 bp in length
*	55535 55634: gap of 100 bp
*	55635 60136: contig of 4502 bp in length
*	60137 60236: gap of 100 bp
*	60237 65623: contig of 5387 bp in length
*	65624 65723: gap of 100 bp
*	65724 70102: contig of 4379 bp in length
*	70103 70202: gap of 100 bp
*	70203 71387: contig of 1185 bp in length
*	71388 71487: gap of 100 bp
*	71488 73835: contig of 2352 bp in length
*	73840 73939: gap of 100 bp
*	73940 77366: contig of 3427 bp in length
*	77367 77466: gap of 100 bp
*	77467 80065: contig of 2599 bp in length
*	80066 80165: gap of 100 bp
*	80166 83825: contig of 3660 bp in length
*	83826 83925: gap of 100 bp
*	83926 86778: gap of 100 bp
*	86778 86877: contig of 2852 bp in length
*	86878 869783: contig of 2906 bp in length
*	869784 89883: gap of 100 bp
*	89884 92751: contig of 2868 bp in length
*	92752 92851: gap of 100 bp
*	92852 95068: contig of 2217 bp in length
*	95069 95168: gap of 100 bp
*	95169 96739: contig of 1571 bp in length
*	96740 96839: gap of 100 bp
*	96840 97912: contig of 1073 bp in length
*	97913 98012: gap of 100 bp
*	98013 99744: contig of 1732 bp in length
*	99745 99844: gap of 100 bp
*	99845 101153: contig of 1309 bp in length
*	101154 101253: gap of 100 bp
*	101254 102868: contig of 1615 bp in length
*	102869 102968: gap of 100 bp
*	102969 104921: contig of 1553 bp in length
*	104922 105021: gap of 100 bp
*	105022 106500: contig of 1479 bp in length
*	106501 106600: gap of 100 bp
*	106601 108529: contig of 1929 bp in length
*	108530 108629: gap of 100 bp
*	108630 110841: contig of 2212 bp in length
*	110842 110941: gap of 100 bp
*	110942 112367: contig of 1426 bp in length
*	112368 112467: gap of 100 bp
*	112468 113665: contig of 1198 bp in length
*	113666 113765: gap of 100 bp
*	113766 114574: contig of 1209 bp in length
*	114575 115074: gap of 100 bp
*	115075 116273: contig of 1199 bp in length.
FEATURES	Location/Qualifiers
SOURCE	1. 116273
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="11"
	/map="11q13"
	/clone="XPac-662A14"
	1. 38138
misc_feature	/note="assembly_fragment"
	38239..44698
misc_feature	/note="assembly_fragment"
	44799..49514
misc_feature	/note="assembly_fragment"
	49615..55534
misc_feature	/note="assembly_fragment"
	55635..60136

Query	Best Local Similarity	Score	DB 2:	Length	Matches	587, Conservative	0; Mismatches	1; Indels	0; Gaps
misc_feature	/note="assembly_fragment"	60237.	.65623						
misc_feature	/note="assembly_fragment"	65724.	.70102						
misc_feature	/note="assembly_fragment"	70203.	.71387						
misc_feature	/note="assembly_fragment"	71488.	.73839						
misc_feature	/note="assembly_fragment"	73940.	.77366						
misc_feature	/note="assembly_fragment"	77467.	.80065						
misc_feature	/note="assembly_fragment"	80166.	.83825						
misc_feature	/note="assembly_fragment"	83926.	.86777						
misc_feature	/note="assembly_fragment"	86878.	.89783						
misc_feature	/note="assembly_fragment"	89884.	.92751						
misc_feature	/note="assembly_fragment"	92852.	.95068						
misc_feature	/note="assembly_fragment"	95169.	.96739						
misc_feature	/note="assembly_fragment"	96840.	.97912						
misc_feature	/note="assembly_fragment"	98013.	.99744						
misc_feature	/note="assembly_fragment"	99845.	.101153						
misc_feature	/note="assembly_fragment"	101254.	.102868						
misc_feature	/note="assembly_fragment"	102969.	.104921						
misc_feature	/note="assembly_fragment"	105022.	.106500						
misc_feature	/note="assembly_fragment"	106601.	.108529						
Query Match		34.48;	Score 586.4;	DB 2:	Length 116273;				
Best Local Similarity		99.88;	Pred. No. 8.2e-105;						
Matches 587, Conservative		0; Mismatches	1;	Indels	0;	Gaps	0;		
QY 1	aagaactggaagacgtctctccactacgacgtgatcgaactggaatccgggccaatg	60							
Db 111636	AAGACGAGGAAAGCTTCCACTACGCGTGATGACAGTGTGAGCGGCCCATG	111695							
QY 61	atggaactggaagcgagggatcaatacagacttaattgaaatataaaggaatatlt	120							
Db 111696	ATGGAATTCGATCCGAGATCAATCAAGCTTAATTTAAATTAATAAAGAAATATTT	111755							
QY 121	tctcccttgaacttatctccgtaaagacatgtgctcctctcttgagggtacggttca	180							
Db 111756	TCTCCCTTTGAACCTTATCTCCGTAAGGACATGTGCTCTCTTTGGGGTCAAGTTC	111815							
QY 181	caatcaatgaccttgaagagctcttgaaatgaagtgtgaagccttggagatttcaagt	240							
Db 111816	CAATCAATGACCTTTGAGAGAGCTTTGATCAAGTTGAGGCGCTTGGAGATTCAAG	111875							
QY 241	cttcaatcgtgttttatcttccctcctcctcaatgttatcctaaccctacatactcgt	300							
Db 111876	CTTCAATCTGGTTTATCTCTCCCTCTCTCANTATTATTAATCCCATATATCTCTAG	111935							
QY 301	aacttgcgtgagccatctcgtgtcatctgtgtgctgggtccacatgctggacaataact	360							
Db 111936	AACCTTGCCTGACGACCATTCCTGGTATGCTGCTGGGCTGCACATGCTGGAACATTA	111995							
QY 361	ggatcgtglaatgaactggaatcctcagatgaagatccctcttgagaatcctatacca	420							
Db 111996	GGATCTGGTATGAACCTGGAATCCTAGTGAAGATCCCTCTTGGAACCTTATTC	112055							
QY 421	ctaactcaaatctggaagcgaagatgtcgtcgtttgtcatcccaagtgtgcaagt	480							

Page 13

Best Local Similarity	76.6%	Prod. No. 3.1e-68	Indels	0	Gaps	0
Matches	490; Conservative	0; Mismatches	150;			
QY 64	agaagatgatgagagacaataaagcttaattgatttaataaagaatattct	123				
Db 146	AGGAGCTGGGGTCAGGATCAAAACACATTTATGTGTGCTTAGGAAAGAAACATTTTCC	205				
QY 124	cccttggaactatctccglaagccatttgctccctcttgggggtcaagtggtcaaa	183				
Db 206	CTCTTTAACTCTCTGAGTACAGTCATTTTGCCCTCTGAGGATCAACTGTCAAC	265				
QY 184	tcaatggcctttagagagctcttgtagtaaatgttagagccttggagatttcagatgct	243				
Db 266	TCAATGGCTTTCAGGACCTCTGGGTCTACCGCTGGTGAACCTGGAGATTCACATCTTT	325				
QY 244	catctggttttatctctccctctccatgattatcauccctcatatctagttagaac	303				
Db 326	CAGACTGTTTCTCTCAATCTTTGGCTGTGTGCTACATCTTATGCTGAGAGAC	385				
QY 304	ttgtctgcaagcaatctcgtgtaatcgttctgtgtgtccaatgctggaacaatacttga	363				
Db 386	TTCACGTGCTTCAATACCTGGSCATGCTGCTGGGTCCACATCTCTGGACATACCTGTTC	445				
QY 364	tcctggaatgaagacgtgaatctcaatgaagaagcctctgttagaattccatccacta	423				
Db 446	TCTGACAAATGACACCTGGGACCTCCAGCCACAAATGACATCTTGAATTCCTCACTCG	505				
QY 424	gaactcaatctgagggcagagaagaatgctgtctgttgcatttccatcccaatgtaggccttct	483				
Db 506	GACTCAAACTGTGGGCAAGAGAGTGTGTGCTTTGTTCCTCAGTGGACGCTCTT	565				
QY 484	caactgaatggagactatccacagcacaaagttagagagacagaaaccccttgttagatgac	543				
Db 566	CACTGAAATGGAGCTTCCCCACACAAATGACGACAGACATGAGACGCTGTGTGAGAGGC	625				
QY 544	tgggtatataatcaaaagctactcccttgacacattgtgactaaagttagagactgtatgt	603				
Db 626	TGGGTATGACAAATCTCTTCTCATCCACCATGCTGACTGAGTGGAGTGGTATGT	685				
QY 604	gattatcagcatgaaatcagtgagtgtaacttctactcttgcactgtaggaatgctgtgtgga	663				
Db 686	GACTGTCATCTCATCTCACTTCAGTGGCTAAATTTGTATTCATGAGCTTGCAATGATGGTGGGA	745				
QY 664	ggcatcatagatgtgccaatgctcagacagatgagctgtgtgtg	703				
Db 746	GGCATCTTAGGGCGTCAATTATCAACAGTAGCCGAGTGG	785				
RESULT 11						
AX179749	AX179749	1986 bp	DNA	linear	PF 06-AUG-2001	
LOCUS	Sequence	37	from Patent	W00146258.		
DEFINITION	AX179749					
ACCESSION	AX179749.1	GI:15132113				
VERSION						
KEYWORDS	human.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
	1 (bases 1 to 1986)					
	Baughn, M.R., Burford, N., Au-Young, J., Lu, D.A., Yang, J., Reddy, R.,					
	Lai, P., Hillman, J., Azimzai, Y., Yue, H., Nguyen, D.B., Yao, M.G.,					
	Gandhi, A.R., Yang, T. and Khan, F. A.					
	Transporters and ion channels					
	Patent: WO 0146258-A 37 28-JUN-2001;					
TITLE	Incyte Genomics, Inc. (US)					
JOURNAL	Location/Qualifiers					
	1..1986					
FEATURES	Organism="Homo sapiens"					
Source	/db_xref="taxon:9606"					
	/note="Incyte ID No: 3358383CB1"					
BASE COUNT	505 a 496 c 457 g 528 t					
ORIGIN						


```

RESULT 13
LOCUS AX074149 1977 bp DNA linear PAT 06-FEB-2001
DEFINITION Sequence 5 from Patent WO0104283.
ACCESSION AX074149
VERSION AX074149.1 GI:12710361
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Sun, W.
TITLE Organic anion transporter genes and proteins
JOURNAL Patent: WO 0104283-A 5 18-JAN-2001;
Metabasis Therapeutics Inc. (US)
FEATURES
source 1..1977
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 502 a 491 c 450 g 534 t
ORIGIN
Query Match 22.2% Score 378; DB 6; Length 1977;
Best Local Similarity 75.9%; Pred. No. 5.9e-64;
Matches 481; Conservative 0; Mismatches 150; Indels 3; Gaps 1;
OY 64 agaaatgatgcgaggaatcaatacaagaatctgaatgaataaagaataatttct 123
DB 113 AGAGAGTGGGGGAGGAGCAACAAACATTAGTGTAGACTTGAAGAAAGAAACATTTC 172
OY 124 ccccttgaaatattccgtaagaacattgtcctcctcttgagggtgaacgtgtcaaa 183
DB 173 CTCCTTGAACCTCTGTGATACGATCTTTCCTTCTACTTGAAGATCAACTGTCAAC 232
OY 184 tcaatgacctgaagagctcttgatgaagcttgaggagcttgaggatctgaatgctt 243
DB 233 TCAATGGCTTTCAAGAGCTCTCTGGGTCAAGCTGTGACCTTGGAGATTCAGATCTT 292
OY 244 catctggtttatctctcctcctcctcctcctcctcctcctcctcctcctcctcct 303
DB 293 CAGACTGTTTCTCTCAATCTTGTGCTGATACATACCTTCATTATTGCTGGAGAAC 352
OY 304 ttgtgtgagcattccctcctcctcctcctcctcctcctcctcctcctcctcctcct 363
DB 353 TTCCTGATTCATACCTGGCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 412
OY 364 tctgttaataaagctgaatcctcctcctcctcctcctcctcctcctcctcctcctcct 423
DB 413 TCTGACATGACACCTGGGGCTCTAGCCCAAGATGACATCTTGAGATCTCCACTTG 472
OY 424 gactcaaatctgagagcagagaagtgtcgtcgtctgtcctcctcctcctcctcctcct 483
DB 473 GACTCAAAACATGAGGCCAGAGAGAGTGTGCTGCTTGTTCATCTCAGAGGAGCTCT 532
OY 484 caccatgagtgagactcaccacagcagagtgaggagcagacagacccctgtgtgatggc 543
DB 533 CACCTGAATGGAGCTTCCCAACACAAAGTGACGACAGATGAGACCTCTGTGTGATGGC 592
OY 544 tgggtatatagaataagctactcctcctcctcctcctcctcctcctcctcctcctcct 600
DB 593 TGGGTGATGACAGAAATCTCTTCTCATCCACATCGGTGAAGCTGAAGTGGATCTGGA 652
OY 601 tgtgattatcagctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 660
DB 653 TGTGACTCTCATCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 712
OY 661 ggaagcattcagatgagcagatgctcagacaggt 694
DB 713 GGAGGCACTTAGCGGTTCATTTATCAGACAGGT 746

```

```

RESULT 14
LOCUS AP003420/c 160945 bp DNA linear PRI 28-AUG-2001
DEFINITION Homo sapiens genomic DNA, chromosome 11q, clone:RP11-614K12,
complete sequence.
ACCESSION AP003420
VERSION AP003420.2 GI:15320509
KEYWORDS
SOURCE HTG.
ORGANISM Homo sapiens
Homo sapiens DNA, clone:RP11-614K12.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Published Only in Database (2001) In press
JOURNAL 2 (bases 1 to 160945)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
Direct Submission
Submitted (16-MAR-2001) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suenho-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
On Aug 27, 2001 this sequence version replaced gi:13383333.
FEATURES
source 1..160945
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-614K12"
BASE COUNT 46743 a 31738 c 32047 g 50417 t
ORIGIN
Query Match 20.0% Score 340.8; DB 9; Length 160945;
Best Local Similarity 76.6%; Pred. No. 6e-57;
Matches 444; Conservative 0; Mismatches 132; Indels 4; Gaps 2;
OY 11 gaagctcttccacacacagctgctatgacacgtgagctccggccacatgagaaatt 70
DB 83641 GAGGAATTTGCTTCCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 83582
OY 71 gatgcagagata--atacaagcttaattgaattgaataaagaagaattttctcctt 128
DB 83581 GGGGTGAGATCAACACACACACTTGTGATTTAGGAAGAAATATTTTCT--CTT 83524
OY 129 tgaacttatccgtaaaagcattgtcctcctcctcctcctcctcctcctcctcctcctcct 188
DB 83523 TGAACCTATCTGGATACACATCTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 83464
OY 189 ggcctttggagagcctctgagtcagcttgaggagccttgaggagattcgaatgctcatct 248
DB 83463 GGCCTTTCAGACCTCTCAGATCAAGTTGAGAGCCCTGGGGAATTCAGATCTTCAGAT 83404
OY 249 gtttttattctctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 308
DB 83403 GGTTCCTTATTAATGTTCAAGCTCATAGTATATCAATCAACCTGAGAGACTTGCC 83344
OY 309 tgcagccattcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 368
DB 83343 ACCATTCATACCTTGATCAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 83284
OY 369 taatgaactgaatccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 428
DB 83283 CATATACCTGGAGACCTTACGACAGAGATCCCTCTGGAATCTCATCTCCATTGGACTC 83224
OY 429 aaatctgagccagagaagtgtcgtcgtcttgatccatcccgatggcagctctcctcctcct 488

```

Db 83223 AATCTGAGCCAGAGAGTGTCTGCTTTGTCATCCCACTGAGAGCTCATTCATCT 83164

Qy 489 gaatggagctctccacagacagatgagcagacagacacccctgtgtgatgctgggt 548
 Db 83163 GAATGGAGCTTCCCAACAGATGAGCCAGATACAGAGAGCCCTGTGTGATGTGGGT 83104

Qy 549 atatatcaagctactctccctgcacattgtgactaa 588
 Db 83103 AATATGACCAAGCTCTTCCCTTCACCATTTGAGTAAAG 83064

RESULT 15
 AC090778 164809 bp DNA linear HTG 27-JAN-2002
 LOCUS Homo sapiens chromosome 11 clone CTD-3110P2 map 11, WORKING DRAFT
 DEFINITION AC090778.2 GI:1436491
 AC090778.2 GI:1436491
 VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 KEYWORDS human.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 164809)
 AUTHORS Biren, B., Linton, L., Nusbaum, C. and Lander, E.
 TITLE Homo sapiens chromosome 11, clone CTD-3110P2
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 164809)
 AUTHORS Biren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S., Batra, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Choquel, Y., Colangelo, M., Collins, S., Collamore, A., Cooke, P., DeArle, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karatas, A., Lacroque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Margulis, N., Mathews, C., McCarthy, M., McEwan, P., McKernan, K., McSheeters, R., Meldrum, J., Menes, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phukhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Sougniez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Travers, M., Travis, N., Trifoglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 TITLE Direct Submission
 JOURNAL Submitted (10-MAR-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 COMMENT On Jun 11, 2001 this sequence version replaced gi:13270647.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

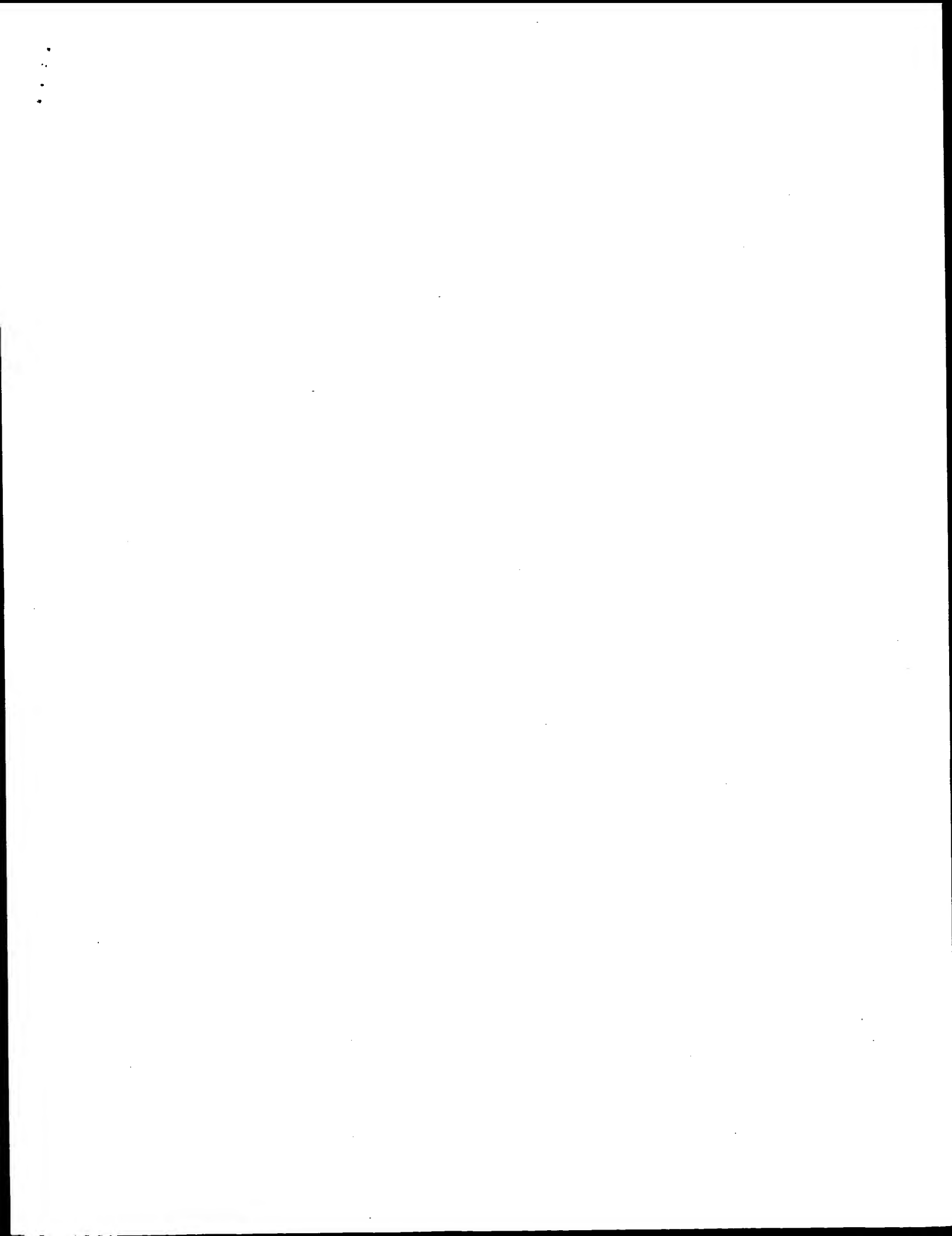
 Genome Center
 Center: Whitehead Institute/MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu

 Project Information
 Center project name: L12735
 Center clone name: 3110.P.2

 Summary Statistics
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 157859 bases at least Q40
 Consensus quality: 161528 bases at least Q30
 Consensus quality: 162934 bases at least Q20
 Insert size: 172000; agarose-IP
 Insert size: 163409; sum-of-contigs

Quality coverage: 5.0 in Q20 bases; agarose-IP
 Quality coverage: 5.2 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 15 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 1069: contig of 1069 bp in length
 * 1070 1169: gap of 100 bp
 * 1170 2379: contig of 1210 bp in length
 * 2380 2479: gap of 100 bp
 * 2480 3440: contig of 961 bp in length
 * 3441 3540: gap of 100 bp
 * 3541 4880: contig of 1340 bp in length
 * 4881 4980: gap of 100 bp
 * 4981 6768: contig of 1788 bp in length
 * 6769 6868: gap of 100 bp
 * 6869 8318: contig of 1450 bp in length
 * 8319 8418: gap of 100 bp
 * 8419 10046: contig of 1628 bp in length
 * 10047 10146: gap of 100 bp
 * 10147 12455: contig of 2309 bp in length
 * 12456 12555: gap of 100 bp
 * 12556 15164: contig of 2609 bp in length
 * 15165 15264: gap of 100 bp
 * 15265 16648: contig of 1384 bp in length
 * 16649 16748: gap of 100 bp
 * 16749 29429: contig of 12681 bp in length
 * 29430 29529: gap of 100 bp
 * 29530 41226: contig of 11697 bp in length
 * 41227 41326: gap of 100 bp
 * 41327 57470: contig of 16044 bp in length
 * 57471 102968: contig of 45498 bp in length
 * 102969 103068: gap of 100 bp
 * 103069 164809: contig of 61741 bp in length.
 *
 * Location/Qualifiers
 * 1. 164809
 * /organism="Homo sapiens"
 * /db_xref="taxon:9606"
 * /chromosome="11"
 * /map="11"
 * /clone="CTD-3110P2"
 * /clone_11b="CTD Human BAC"
 * 1. 1069
 * /note="assembly-fragment"
 * 1170. 2379
 * /note="assembly-fragment"
 * 2480. 3440
 * /note="assembly-fragment"
 * 3541. 4880
 * /note="assembly-fragment"
 * 4981. 6768
 * /note="assembly-fragment"
 * 6869. 8318
 * /note="assembly-fragment"
 * 8419. 10046
 * /note="assembly-fragment"
 * 10147. 12455
 * /note="assembly-fragment"
 * 12556. 15164
 * /note="assembly-fragment"
 * 15265. 16648
 * /note="assembly-fragment"
 * 16749. 29429
 * /note="assembly-fragment"
 * 29530. 41226
 * /note="assembly-fragment"
 * 41327. 57370
 * /note="assembly-fragment"



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 1, 2002, 12:39:46 ; Search time 3305.87 Seconds
(without alignments)
5248.165 Million cell updates/sec

Title: US-09-674-235-10

Sequence: 1 atggccttgaggagctctt.....acacaaagtgcacaaattt 804

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenDbml:*

1: gb_da:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_to:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_inv:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Match	Length	DB	ID	Description
------------	-------	-------	--------	----	----	-------------

1	508	63.2	2684	6	AX074150	AX074150 Sequence
2	402	50.0	160945	2	AP003420	AP003420 Homo sapi
3	402	50.0	166848	2	AP002367	AP002367 Homo sapi
4	402	50.0	169776	9	AP001880	AP001880 Homo sapi
5	402	50.0	170682	2	AC073256	AC073256 Homo sapi
6	400.4	49.8	116273	2	AP000596	AP000596 Homo sapi
7	336.2	41.7	1566	9	BC022378	BC022378 Homo sapi
8	335.2	41.7	1846	9	AB062418	AB062418 Homo sapi
9	335.2	41.7	1986	6	AX179749	AX179749 Sequence
10	314.2	39.1	1977	6	AX074149	AX074149 Sequence
11	300	37.3	2597	10	RN018	RN018 Rattus norv
12	293.6	36.5	1995	10	AB051836	AB051836 Rattus nor
13	293.6	36.5	2253	9	AK057654	AK057654 Homo sapi
14	290.4	36.1	1964	10	BC021384	BC021384 Mus muscu
15	288.8	35.9	1911	10	RN013285	RN013285 Rattus no
16	277.6	34.5	2016	10	BC014805	BC014805 Mus muscu
17	277.6	34.5	2878	10	MMU132857	MMU132857 Mus muscu
18	275.6	34.3	116273	2	AP000596	AP000596 Homo sapi
19	275.6	34.3	160945	2	AP003420	AP003420 Homo sapi
20	275.6	34.3	164809	2	AC090778	AC090778 Homo sapi
21	269.6	33.5	169776	9	AP001880	AP001880 Homo sapi
22	269.6	33.5	2079	10	RN013285	RN013285 Rattus no
23	261.2	32.5	166848	2	AP002367	AP002367 Homo sapi
24	250	31.1	179934	9	AP000810	AP000810 Homo sapi
25	250	31.1	203598	2	AC090250	AC090250 Homo sapi
26	250	31.1	210723	2	AP001858	AP001858 Homo sapi
27	241.6	30.0	181234	2	AC103148	AC103148 Rattus no
28	241.6	30.0	185404	2	AC025409	AC025409 Homo sapi
29	238.8	29.7	207064	2	AC106680	AC106680 Rattus no
30	211.8	26.0	170682	2	AC073256	AC073256 Homo sapi
31	208.8	26.0	1737	6	AX191489	AX191489 Sequence
32	208.8	26.0	2455	6	AX230568	AX230568 Sequence
33	208.8	26.0	2665	6	AX191499	AX191499 Sequence
34	191.2	23.8	1779	10	AB005451	AB005451 Mus muscu
35	181	22.5	2516	9	AK055737	AK055737 Homo sapi
36	180.8	22.5	1653	6	AX191610	AX191610 Sequence
37	180.8	22.5	2047	6	AX230572	AX230572 Sequence
38	180.8	22.5	2194	6	AX191620	AX191620 Sequence
39	180.8	22.5	2210	6	E54995	E54995 Placental o
40	180.8	22.5	2210	9	AB026116	AB026116 Homo sapi
41	179.6	22.3	3258	9	AB050269	AB050269 Homo sapi
42	174.8	21.7	79141	2	AC099123	AC099123 Rattus no
43	160.4	20.0	119151	2	AP000655	AP000655 Homo sapi
44	160.4	20.0	120628	9	AC012153	AC012153 Homo sapi
45	160.4	20.0	153169	2	AP001092	AP001092 Homo sapi

ALIGNMENTS

RESULT 1	AX074150	AX074150	2684 bp	DNA	linear	PAT 06-FEB-2001
LOCUS	AX074150	Sequence	6 from Patent WO0104283.			
DEFINITION	AX074150					
ACCESSION	AX074150					
VERSION	AX074150.1	GI:12710362				
KEYWORDS						
SOURCE		human.				
ORGANISM		Homo sapiens				
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1	(bases 1 to 2684)				
AUTHORS	Sun, W.					
TITLE	Organic anion transporter genes and proteins					
JOURNAL	Patent: WO 0104283-A 6 18-JAN-2001;					
	Metabasis Therapeutics Inc. (US)					
FEATURES	Location/Qualifiers					
	1..2684					
	/organism="Homo sapiens"					
	/db_xref="taxon:9606"					
BASE COUNT	800 a	533 c	541 g	810 t		
ORIGIN						

Query Match 63.2%; Score 508; DB 6; Length 2684;
 Best Local Similarity 100.0%; Pred. No. 2.4e-136;
 Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 atggcctttgagagagctcttgatcaagttgagagccttgagagattcagatgctcat 60
 |||||||
 Db 200 ATGGCTTTGAGAGAGCTCTTGATCAAGTTGAGAGCCTTGAGAGATTTCAGATGCTTCAT 259
 OY 61 ctggtttattcttccttcctcatgattatattccctatctactgctcagaagactt 120
 |||||||
 Db 260 CTGGTTTATTCTTCTCCCTCTCATGTTATTATCCCTATATACCTGCTAGAGAACTTT 319
 OY 121 gctgagcattcttcgtgcatcgttgctgggtccacatgctgagacaataactgactt 180
 |||||||
 Db 320 GCTGACGACATTCCTGCTCATGCTGCTGGTCCACATGCTGACAAATATATGATCT 379
 OY 181 ggtatgaactcggaaactcctcaagtgaagatgcctcttgagaactctatccactagac 240
 |||||||
 Db 380 GGTATATGAAGACTGGAACTCTCAGTGAAGATGCCCTCTTGAGAACTCTATCCACTAGAC 439
 OY 241 tcaaatctgagccagagaagtgctgctcttgctaccccaatgagagctcttcac 300
 |||||||
 Db 440 TCAAACTGAGGCCAGAGAGAGTGTCTCCCTTTGCTCATCCCGAGTGGAGCTTTCTCAC 499
 OY 301 ctgaatggagactatcacagacaaagttagagagacagaaacctgtgtgagtgctg 360
 |||||||
 Db 500 CTGAATGGGACTATCCACAGCAGCAAGTGAAGCAGACAGAACCCCTGTGGATGGCTGG 559
 OY 361 gtaattgaaagacttactccttccttcgacattgagcaaatggagagctgtatgtat 420
 |||||||
 Db 560 GTATATGATCAAGACTACTCTCTTGACCAATGTGATCAAGTGGGACCTGTGATGTGAT 619
 OY 421 tatcagtcactgaatcagtggttcaatcctactcttactggaagtgtgtgagagac 480
 |||||||
 Db 620 TATCAGTCACTGAATCAAGTGTGTCATCTTCTGACACGATGCTGTGGTGGAGAGC 679
 OY 481 atcataagtgccatgcttcagacaggt 508
 |||||||
 Db 680 ATCATAGGTGGCCATGTCTCAGACAGGT 707

RESULT 2
 AP003420 160945 bp DNA linear PRI 28-AUG-2001
 LOCUS Homo sapiens genomic DNA, chromosome 11q, clone:RP11-614K12,
 DEFINITION complete sequence.
 ACCESSION AP003420
 VERSION AP003420.2 GI:15320509
 KEYWORDS HTG.
 SOURCE Homo sapiens DNA, clone:RP11-614K12.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 TITLE Homo sapiens genomic DNA
 JOURNAL Published Only in Database (2001) In press
 REFERENCE 2 (bases 1 to 160945)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (16-MAR-2001) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suenho-cho, Tsukuba, Ibaraki, Japan (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/),
 Tel:81-45-503-9111, Fax:81-45-503-9170
 COMMENT On Aug 27, 2001 this sequence version replaced gi:15383333.
 FEATURES
 source location/Qualifiers
 1..160945
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

/chromosome="11"
 /map="11q"
 /clone="RP11-614K12"
 BASE COUNT 46743 a 31738 c 32047 g 50417 t
 ORIGIN

Query Match 50.0%; Score 402; DB 9; Length 160945;
 Best Local Similarity 100.0%; Pred. No. 1.5e-105;
 Matches 402; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 atggcctttgagagagctcttgatcaagttgagagccttgagagattcagatgctcat 60
 |||||||
 Db 143970 ATGGCTTTGAGAGAGCTCTTGATCAAGTTGAGAGCCTTGAGAGATTTCAGATGCTTCAT 144029
 OY 61 ctggtttattcttccttcctcatgattatattccctatctactgctcagaagactt 120
 |||||||
 Db 144030 CTGGTTTATTCTTCTCCCTCTCATGTTATTATCCCTATATACCTGCTAGAGAACTTT 144089
 OY 121 gctgagcattcttcgtgcatcgttgctgggtccacatgctgagacaataactgactt 180
 |||||||
 Db 144090 GCTGACGACATTCCTGCTCATGCTGCTGGTCCACATGCTGACAAATATATGATCTG 144149
 OY 181 ggtatgaactcggaaactcctcaagtgaagatgcctcttgagaactctatccactagac 240
 |||||||
 Db 144150 GGTATATGAAGACTGGAACTCTCAGTGAAGATGCCCTCTTGAGAACTCTATCCACTAGAC 144209
 OY 241 tcaaatctgagccagagaagtgctgctcttgctaccccaatgagagctcttcac 300
 |||||||
 Db 144210 TCAAACTGAGGCCAGAGAGAGTGTCTCCCTTTGCTCATCCCGAGTGGAGCTTTCTCAC 144269
 OY 301 ctgaatggagactatcacagacaaagttagagagacagaaacctgtgtgagtgctg 360
 |||||||
 Db 144270 CTGAATGGGACTATCCACAGCAGCAAGTGAAGCAGACAGAACCCCTGTGGATGGCTGG 144329
 OY 361 gtaattgaaagacttactccttccttcgacattgagcaaatggagagctgtatgtat 402
 |||||||
 Db 144330 GTATATGATCAAGACTACTCTCTTGACCAATGTGATGATCAAG 144371

RESULT 3
 AP002367 166848 bp DNA linear HTG 31-MAY-2000
 LOCUS Homo sapiens chromosome 11 clone RP11-151E18 map 11q, WORKING DRAFT
 DEFINITION SEQUENCE, 9 unordered pieces.
 ACCESSION AP002367
 VERSION AP002367.1 GI:8131631
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens DNA, clone:RP11-151E18.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 TITLE Homo sapiens 166,848 genomic DNA of 11q
 JOURNAL Published Only in Database (2000) In press
 REFERENCE 2 (bases 1 to 166848)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 TITLE Direct Submission
 JOURNAL Submitted (29-MAY-2000) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
 Japan (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/),
 Tel:81-42-778-9923, Fax:81-42-778-9924

----- Genome Center
 Center: RIKEN Genomic Sciences Center(GSC)
 Center code: RIKEN
 Web site: http://hgp.gsc.riken.go.jp/
 Contact: hattori@gsc.riken.go.jp
 ----- Project Information


```

/note="assembly_fragment clone_end:T7 vector_side:left"
misc_feature      162606..165410
/note="assembly_fragment clone_end:SP6 vector_side:left"
misc_feature      163511..166848
/note="assembly_fragment"
BASE_COUNT      55669 a 32391 c 31923 g 46064 t 801 others
ORIGIN

```

Query Match	50.0%	Score 402	DB 2	Length 166848
Best Local Similarity	100.0%	Pred. No. 1.5e-105	Mismatches 0	Indels 0
Matches 402	Conservative	0		
Qy 1	atggccttgaagagactcttgagtaagaattggaagccttggagaatttcagatgcttcat	60		
Db 161305	ATGGCCTTGGAGAGCTCTTGAGTCAAGTTGGAGGCGCTTGGAGATTTCAGATCTTCAT	161364		
Qy 61	ctggctttatctctccctctcctaigtatttaattccctcatatctgctagaaacttt	120		
Db 161365	CTGGCTTTATCTCCCTCTCCTCAATGTTATTTATCCCTCATATCTGCTTGSAGAACTT	161424		
Qy 121	gcttgagccatctctggtcaatgcgttgcgttgacacatgctggaataatactgagatc	180		
Db 161425	GCTGAGCCATCTCCGTCATCGTTCCTGGCTGCACATGCTGACAAATATCTGAGTCT	161484		
Qy 181	ggtgaatgaactggaatcctcagtgaaagatgcccctcttgagaatctctatccactagac	240		
Db 161485	GGTGAATGAACTCGAATCCTCTAGTGAATATCCCTCTTGGAATCTCTATCCACACTGAC	161544		
Qy 241	tcaatctgagggcagagaagaatggtctgctgttgcataccccaagtggagactcttcac	300		
Db 161545	TCAAATCTGGAGGCAAGAAAGTGTGCTGTTCATCCCGAGTGCGACTTCTTCAC	161604		
Qy 301	ctgaatgagacataccacagacacaaagtgaaggagagacaaagaaccttgtgagtgctag	360		
Db 161605	CTGAATGGAGCATCCACAGACAAAGTGAAGGAGACACAGAACCTCTGTGTGATGCGCTTG	161664		
Qy 361	gataatgataaagactacttcccttgacacattgagactag 402			
Db 161665	GTAATATGATCAAAAGCTACTTCCCTTGACACATTGGAGCTAAG 161706			
RESULT 4				
AP001880	169776 bp DNA linear PRI 26-APR-2001			
LOCUS				
DEFINITION	Homo sapiens genomic DNA, chromosome 11q, clone:CMD-3110P2,			
ACCESSION	AP001880			
VERSION	AP001880			
KEYWORDS	HTG.			
SOURCE	Homo sapiens DNA, clone:CMD-3110P2.			
ORGANISM	Homo sapiens			
REFERENCE	Homo sapiens			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.			
TITLE	1 (sites)			
JOURNAL	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,			
REFERENCE	Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.			
AUTHORS	Homo sapiens genomic DNA			
	Published Only in Database (2000) In press			
	2 (bases 1 to 169776)			
	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,			
	Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.			
	Direct Submission			
	Submitted (24-APR-2000) Masahira Hattori, The Institute of Physical			
	and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Japan			
	1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan			
	(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/)			
	Tel:81-45-503-9111, Fax:81-45-503-9170)			
	On Apr 26, 2001 this sequence version replaced gi:10716820.			
COMMENT	Location/Qualifiers			
FEATURES	1..169776			
SOURCE	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			

106601 108529 contig of 1929 bp in length
 108630 110841 contig of 2212 bp in length
 110942 112367 contig of 1426 bp in length
 112468 113665 contig of 1198 bp in length
 113766 114974 contig of 1209 bp in length
 115075 116273 contig of 1199 bp in length

Sequence updated (26-May-2000).

* NOTE: this is a 'working draft' sequence. It currently
 * consists of 29 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 38138: contig of 38138 bp in length
 38139 38238: gap of 100 bp
 38239 44698: contig of 6460 bp in length
 44699 44798: gap of 100 bp
 44799 49514: contig of 4716 bp in length
 49515 49614: gap of 100 bp
 49615 55534: contig of 5920 bp in length
 55535 55634: gap of 100 bp
 55635 60136: contig of 4502 bp in length
 60137 60236: gap of 100 bp
 60237 65623: contig of 5387 bp in length
 65624 65723: gap of 100 bp
 65724 70102: contig of 4379 bp in length
 70103 70202: gap of 100 bp
 70203 71387: contig of 1185 bp in length
 71388 71487: gap of 100 bp
 71488 73839: contig of 2352 bp in length
 73840 73939: gap of 100 bp
 73940 77366: contig of 3427 bp in length
 77367 77466: gap of 100 bp
 77467 80065: contig of 2599 bp in length
 80066 80165: gap of 100 bp
 80166 83825: contig of 3660 bp in length
 83826 83925: gap of 100 bp
 83926 86777: contig of 2852 bp in length
 86778 86877: gap of 100 bp
 86878 89783: contig of 2906 bp in length
 89784 89883: gap of 100 bp
 89884 92751: contig of 2868 bp in length
 92752 92851: gap of 100 bp
 92852 95068: contig of 2217 bp in length
 95069 95168: gap of 100 bp
 95169 96739: contig of 1571 bp in length
 96740 96839: gap of 100 bp
 96840 97912: contig of 1073 bp in length
 97913 98012: gap of 100 bp
 98013 99744: contig of 1732 bp in length
 99745 99844: gap of 100 bp
 99845 101153: contig of 1309 bp in length
 101154 101253: gap of 100 bp
 101254 102868: contig of 1615 bp in length
 102869 102968: gap of 100 bp
 102969 104921: contig of 1953 bp in length
 104922 105021: gap of 100 bp
 105022 106500: contig of 1479 bp in length
 106501 106600: gap of 100 bp
 106601 108529: contig of 1929 bp in length
 108530 108629: gap of 100 bp
 108630 110841: contig of 2212 bp in length
 110842 110941: gap of 100 bp
 110942 112367: contig of 1426 bp in length
 112368 112467: gap of 100 bp
 112468 113665: contig of 1198 bp in length
 113666 113765: gap of 100 bp
 113766 114974: contig of 1209 bp in length
 114975 115074: gap of 100 bp
 115075 116273: contig of 1199 bp in length.

Location/Qualifiers

source
 1..116273
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q13"
 /clone="XPac-662A14"
 1..38138
 /note="assembly_fragment"
 38239..44698
 /note="assembly_fragment"
 44799..49514
 /note="assembly_fragment"
 49615..55534
 /note="assembly_fragment"
 55635..60136
 /note="assembly_fragment"
 60237..65623
 /note="assembly_fragment"
 65724..70102
 /note="assembly_fragment"
 70203..71387
 /note="assembly_fragment"
 71488..73839
 /note="assembly_fragment"
 73940..77366
 /note="assembly_fragment"
 77467..80065
 /note="assembly_fragment"
 80166..83825
 /note="assembly_fragment"
 83926..86777
 /note="assembly_fragment"
 86878..89783
 /note="assembly_fragment"
 89884..92751
 /note="assembly_fragment"
 92852..95068
 /note="assembly_fragment"
 95169..96739
 /note="assembly_fragment"
 96840..97912
 /note="assembly_fragment"
 98013..99744
 /note="assembly_fragment"
 99845..101153
 /note="assembly_fragment"
 101254..102868
 /note="assembly_fragment"
 102969..104921
 /note="assembly_fragment"
 105022..106500
 /note="assembly_fragment"
 106601..108529
 /note="assembly_fragment"

Query Match 49.8%; Score 400.4; DB 2; Length 116273;
 Best Local Similarity 99.8%; Pred.No. 4.3e-105;
 Matches 401; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atggcccttggagggccttgaagtcgaagcttggagggccttggagagatttcaagatgcttcatt 60
 |||||
 Db 111822 ATGGCCTTTGGAGGAGCTTTGAGTCAAGTGGAGGCTTTGGAGATTTCAGATGCTTCA 111881

QY 61 ctggctttattcttcctcctcctcatgttatataccctacatactctagagaacttt 120
 |||||
 Db 111882 CTGGCTTTATTCTTCCTCTCTCATGTATTATACCTCCATATATCTGTAGAGAACTTT 111941

QY 121 gctgcagccattcctcgtgcatcgttgcctgggtccacatgctggaacaataactgatct 180
 |||||
 Db 111942 GCTGCAGCCATTCTCGTGCATCGTTGCTGGGTCACATGCTGGAGCAATATATCTGATCT 112001

QY 181 ggtaatgaactggaatcctcagtggaagatgcctcttgagaattctataccactagac 240
 |||||
 Db 112002 GGTAAATGAACCTGAATCCTCAGTGAAGATGCTCCCTTGAGAAATCTATACCACTAGAC 112061

FEATURES

QY 241 tcaaatctgagccagagaagtgtcgtcttgcattccaccagtgagctcttcac 300
 |||||||
 Db 112062 TCAAAATCTGAGCCAGAGAAAGTGTCGTCCTTGTCATCCAGAGAGCTTCTTAC 112121
 |||||||
 QY 301 ctgaatggagctatccacagcagaatgagcagacagaacccctgtgtgagtcgtg 360
 |||||||
 Db 112122 CTGAATGGAGCTATCCACGACAAAGTAGGACAGACAGAAACCCGTGTGTGATGGCTGG 112181
 |||||||
 QY 361 gctatgacaaagctactccctcgtacattgtgactaag 402
 |||||||
 Db 112182 GTRATGATCAAGCTACTTCCCTTGACCATTTGTGACTTAC 112223
 |||||||
 RESULT 7
 BC022379 1566 bp mRNA linear PRI 04-FEB-2002
 LOCUS Homo sapiens, similar to ustr3, clone MGC:23972 IMAGE:4714598, mRNA,
 DEFINITION complete cds.
 ACCESSION BC022379
 VERSION BC022379.1 GI:18490377
 KEYWORDS MGC.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 1566)
 AUTHORS Strausberg, R.
 JOURNAL Direct Submission
 Submitted (01-FEB-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: CLONTECH
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human
 Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mdickpaxl.stanford.edu
 Dickson, M., Schmitz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 37 Row: d Column: 21
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.
 FEATURES
 source
 1.1566
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:23972 IMAGE:4714598"
 /tissue="type=Liver"
 /clone_lib="NIH_MGC_76"
 /lab_host="DH10B"
 /note="Vector: pDNR-LIB"
 269.985
 /codon_start=1
 /product="Similar to ustr3"
 /protein_id="AAH2379.1"
 /db_xref="GI:18490378"
 /translation="MAFDLGHAGDLMRFQLOTVLSIFAVATYLHMLNENFTAFI
 PGRWCWHLIDNDVTSNDNGALSDALRLISPLDSNMPKRCRFVHPQMLHLN
 GTEPNTSDADMEPCVDGWYDRISFSTLTETEDLDVCDOSLSVAKFVPMAGMNGG
 ILGHSLSDSRNGNTQIRPGHNTIGNVPMWCTIYDGRGFCFCHSLAYPPAGVCVILL
 CDLSLDKLARVCSVAHYIOO"
 CDS
 431 a 375 c 356 g 404 t
 BASE COUNT
 ORIGIN

Query Match 41.8%; Score 336.2; DB 9; Length 1566;
 Best Local Similarity 78.1%; Pred. No. 1,8e-86;
 Matches 404; Conservative 0; Mismatches 113; Indels 0; Gaps 0;
 QY 1 atgagcttggagagactcttgaatgaattggagagcttggagattccagtcctac 60
 |||||||
 Db 269 ATGGCTTTTCAGGACCTCCGCGGTCAAGCTGTGAGACTGTGGAGATTCCAGATCTTCAG 328
 |||||||
 QY 61 ctgttttattctccctcccaatgttaataaccataactcagtagaactt 120
 |||||||
 Db 329 ACTGTTTTCCTCAATCTTTCCTGTTGCTTACATACCTTATTTATGCGGAGAACTTC 388
 |||||||
 QY 121 gctgcagccattcctgtgtacatgctgtgggtccacatgctgagacaataactgactc 180
 |||||||
 Db 389 ACTGATTCATFACCTGGCCATCGCTGCTGGTCCACATCTGAGACATGACACTGTCTCT 448
 |||||||
 QY 181 gctaatgaaactggaatccctcagtgaaatgcccctcttggaatctctatccactagac 240
 |||||||
 Db 449 GACAAATGACACTGGGGCCCTCAGCCAAAGATGACACTCTTGAGAAATCTCCATCCACTGGAC 508
 |||||||
 QY 241 tcaatctgagccagagaagtgtcgtcttgcattccaccagtgagctcttcac 300
 |||||||
 Db 509 TCAAAATCTGAGCCAGAGAAAGTGTCGTCCTTGTCATCCAGAGAGCTTCTTAC 568
 |||||||
 QY 301 ctgaatggagctatccacagcagaatgagcagacagaacccctgtgtgagtcgtg 360
 |||||||
 Db 569 CTGAATGGAGCTATCCACGACAAAGTAGGACAGACAGCCCTGTGTGATGGCTGG 628
 |||||||
 QY 361 gctatgacaaagctactccctcgtacattgtgactaagtgagcctgtagtgcgt 420
 |||||||
 Db 629 GTGATGACAGAAATCTCTTCTCATCCACATGCTGACTGATGAGTGTGATGTGAC 688
 |||||||
 QY 421 tctagctcactgaaatcagtgatctcaattcctactcttgactgagatgctgtggagac 480
 |||||||
 Db 689 TCCCAATGACTGACTTCAGTGGCTAATTTGTATTCAATGCTCGAATGAGTGGAGGAGC 748
 |||||||
 QY 481 atcatagtgccatgctcctcagagagtgctgtg 517
 |||||||
 Db 749 ATCTTAGCGGCTCATTTATCAGACAGTAGCCGAGGTGG 785
 |||||||
 RESULT 8
 AB062418 1846 bp mRNA linear PRI 05-JAN-2002
 LOCUS Homo sapiens ustr3 mRNA for huST3, complete cds.
 DEFINITION
 ACCESSION AB062418
 VERSION AB062418.1 GI:18148872
 KEYWORDS
 SOURCE Homo sapiens tissue_lib:fetal liver cdna to mRNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (sites)
 AUTHORS Takanaga, H., Ohtsuki, S., Hosoya, K., and Terasaki, T.
 TITLE Isolation of novel clone of amphiphilic solute facilitator family
 from human fetal liver
 JOURNAL Unpublished
 2 (bases 1 to 1846)
 Takanaga, H., Ohtsuki, S. and Terasaki, T.
 Direct Submission
 Submitted (26-MAY-2001) Hitomi Takanaga, Tohoku University,
 Graduate School of Pharmaceutical Sciences, Aramaki, Aoba, Sendai,
 Aoba-ku, Miyagi 980-8578, Japan
 (E-mail: takanaga@mail.pharm.tohoku.ac.jp, Tel: 81-22-217-6832,
 Fax: 81-22-217-6886)
 FEATURES
 source
 1.1846
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_lib="fetal liver"
 102.1763
 gene

CDS

/gene="ust3"
102..1763
/gene="ust3"
/codon_start=1
/product="HUST3"
/protein_id="BAB83517.1"
/db_xref="GI:18148873"
/translation="MAFODLGHAGDLMFRLIQLQVFLSIFAVATYLFHLENTAFI
FGHRCVHILDMNDIVSDNDTALSDALIRISIPIDSNMRPEKCRREHPWOMLIHN
GTFPNTSDADMEPCVDGWYDRISFSSTVTEMDLYCSOSITSVAKFYPMGMYGG
ILGHLSPDRGRFRLKWCYLOVAIVGTALAPFTLITCSLHFLSGIAMSLLITIT
MLIEMATHRRDQAGITITGMCPSGLAFMLAGIARIRUMHIIQIVAVPEVIFLTS
SWLESARWLIINNKPEBGLKELRKAHRSKGNKARDTLTLEIKSTMKELAAOK
KPSICEMLMHPICRISILSFTREFNEMAYGLNHYOHLGNVFLQTLFGAVILL
ANCVAPMAIKYNNRRASOMLMLFLAICLAIIFVQEMQTLREVLAITGLGASALAN
TLAFHNGNEVITITIRARMGINATEANAGALAPLMTISVSPPLMIIVGVPEFI
SGFAFLIPGRNKPFLPDTIDQENMRKDRPREKQEDPREVETQF"

BASE COUNT 458 a 472 c 420 g 496 t

ORIGIN

Query Match 41.7%; Score 335.2; DB 9; Length 1846;
Best Local Similarity 78.7%; Pred. No. 3.4e-86;
Matches 400; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 1 atggccttgaggagccttgagtcagttgaggccttgaggatctcagatgctcat 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 102 ATGGCTTTCAGAGACCTCTCGGGTACGCTGGTACCTGGAGATTCCAGATCCTTCAG 161

QY 61 ctggttttattctccctctcctcatgttataatccctcatatgcttagaagactt 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 162 ACTGTTTTCTCTCAATCTTCTGCTGCTACATACCTTCATTTTATGCTGGAGAACTTC 221

QY 121 gctgcagcattcctcgtcatcgttgcgtgctccacatgctgagcaataatactgact 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 222 ACTGATTCATATCCCTGGCCATGCTGCTGGTCCACATCTGGACATGACACTGTCTCT 281

QY 181 ggtaatgaactggaactcctcagtgaaagatgaccttgagaatctctatccactagac 240
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 282 GACATATGACACTGGGCGCTCAGCAAGATGACATCTTGAGATCTCCATCCCACTGGAC 341

QY 241 tcaaatctgaggcagagaaagtgtgtgcttgccttgcacatcccaagtgagcactcttcac 300
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 342 TCATAACATGAGGCCAGAAAGTGTGCTGCTTGTTCATCTCCAGTGGAGAGCTCCCTTCAC 401

QY 301 ctgaatggactatccacagcacaaagttaggcagacacacccctgtgtgtagagctg 360
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 402 CTGAATGGGACCTTCCCAACACAGTACGACACATGGAGCCCTGTGTGATGGCTGG 461

QY 361 gtataatgcaagctacttcccttcgacacattgtgactaagtggagcctgtgtatgat 420
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 462 GTGTATGACAGAAATCTCTTCTCATCCACATCGTACATGAGTGGGATCTGGTATGTGAC 521

QY 421 tatcagtcacatgaaatcagttgattcaattccacttgcactcggaatcgtgtgtagagc 480
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 522 TCTCAATCAGTCACTGCTGAGCTAAATTTGTATTATCATGTGATGATGATGGAGAGC 581

QY 481 atcataagtgagcattgtctcagacagct 508
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 582 ATCCTAGGCGGTCAATTATCAGACAGGT 609

RESULT 9
AX179749 1986 bp DNA linear PAT 06-AUG-2001
LOCUS AX179749
DEFINITION Sequence 37 from Patent WO0146258.
ACCESSION AX179749
VERSION AX179749.1 GI:15132113
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1986)
AUTHORS Baughm, M.R., Butford, N., An-Young, J., Lu, D.A., Yang, J., Reddy, R.,
Lal, P., Hillman, J.L., Azimzal, Y., Yue, H., Nguyen, D.B., Yao, M.G.,
Gandhi, A.R., Yang, Y.T., and Khan, F.A.
TITLE Transporters and ion channels
JOURNAL Patent: WO 0146258-A 37 28-JUN-2001;
Incyte Genomics, Inc. (US)
FEATURES location/Qualifiers
SOURCE 1..1986
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 3358383CB1"

BASE COUNT 505 a 496 c 457 g 528 t

ORIGIN

Query Match 41.7%; Score 335.2; DB 6; Length 1986;
Best Local Similarity 78.7%; Pred. No. 3.4e-86;
Matches 400; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 1 atggccttgaggagccttgagtcagttgaggccttgaggatctcagatgctcat 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 248 ATGGCTTTCAGAGACCTCTCGGGTACGCTGGTACCTGGAGATTCCAGATCCTTCAG 307

QY 61 ctggttttattctccctctcctcatgttataatccctcatatgcttagaagactt 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 308 ACTGTTTTCTCTCAATCTTCTGCTGCTACATACCTTCATTTTATGCTGGAGAACTTC 367

QY 121 gctgcagcattcctcgtcatcgttgcgtggtgcacatgctgagcaataatactgact 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 368 ACTGATTCATATCCCTGGCCATGCTGCTGGTCCACATCTGGACATGACACTGTCTCT 427

QY 181 ggtaatgaactggaactcctcagtgaaagatgaccttgagaatctctatccactagac 240
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 428 GACATATGACACTGGGCGCTCAGCAAGATGACATCTTGAGATCTCCATCCCACTGGAC 487

QY 241 tcaaatctgaggcagagaaagtgtgtgcttgccttgcacatcccaagtgagcactcttcac 300
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 488 TCATAACATGAGGCCAGAAAGTGTGCTGCTTGTTCATCTCCAGTGGAGAGCTCCCTTCAC 547

QY 301 ctgaatggactatccacagcacaaagttaggcagacacacacccctgtgtgtagagctg 360
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 548 CTGAATGGGACCTTCCCAACACAGTACGACACATGGAGCCCTGTGTGATGGCTGG 607

QY 361 gtataatgcaagctacttcccttcgacacattgtgactaagtggagcctgtgtatgat 420
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 608 GTGTATGACAGAAATCTCTTCTCATCCACATCGTACATGAGTGGGATCTGGTATGTGAC 667

QY 421 tatcagtcacatgaaatcagttgattcaattccacttgcactcggaatcgtgtgtagagc 480
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 668 TCTCAATCAGTCACTGCTGAGCTAAATTTGTATTATCATGTGATGATGATGGAGAGC 727

QY 481 atcataagtgagcattgtctcagacagct 508
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 728 ATCCTAGGCGGTCAATTATCAGACAGGT 755

RESULT 10
AX074149 1977 bp DNA linear PAT 06-FEB-2001
LOCUS AX074149
DEFINITION Sequence 5 from Patent WO0104283.
ACCESSION AX074149
VERSION AX074149.1 GI:12710361
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1977)
AUTHORS Sun, W.
TITLE Organic anion transporter genes and proteins
JOURNAL Patent: WO 0104283-A 5 18-JAN-2001;
Metabasis Therapeutics Inc. (US)

FEATURES		Location/Qualifiers	
	source	1..1977	/organism="Homo sapiens" /db_xref="taxon:9606"
BASE COUNT	502 a 491 c 450 g 534 t		
ORIGIN			
Query Match	39.1%; Score 314.2; DB 6; Length 1977;		
Best Local Similarity	77.3%; Pred No. 4.3e-80;		
Matches	395; Conservative 0; Mismatches 113; Indels 3; Gaps 1.		
OY	1 atggcccttgaagagcgtcttgatgcacgcttgtagaggcccttgtagagattcaagtgtcat 60		
Dd	236 ATGGCCTTTCAGACACTCCCTGCGTCACCGCTGTGAACCTGTGAGATTCACATCTCTTGAC 295		
OY	61 ctggttttatcttcctccctctcatalgtatatcaatccctcatatactctagagaactt 120		
Dd	296 ACTGTTTCTCTCAATCTTTGCTGTGTGCATCATACCTTCATTATTCGTGGAACTTTC 355		
OY	121 gctgcagccatctcctgycatcgttgtctcggtgcaccatcgtcgtgacaataatcigtatct 180		
Dd	356 ACTGCATTCATACCTGGCCATCGCTGTGGGCCACATCCCTGGACAAATGACACTGTCTCT 415		
OY	181 ggtaatgaacctggaatcctcaagtgaagaatgccctctcttgagaattctatacccattagac 240		
Dd	416 GACATTCACACTGGGGCCCTCAGCCAAATATGACATCTTGAGATCTTCATCCCACTGGAC 475		
OY	241 tcaaatctgaggccagagaagltgctgctcttgcttcacatccccagtgcaactcttac 300		
Dd	476 TCAAACATGAGGCCAGAGAAGTGTGTGCTTTGTTCATCCCATGGCACGCTCTTACC 535		
OY	301 ctgaatgagactatccaagcacaaagtgaaggacagacacagaaccctgtgtgatgctg 360		
Dd	536 CTGAATGGGACCTTCCCAACAAGTACACGACATGGACATGGACCCCTGTGTGATGGCTGG 595		
OY	361 gtatatgaccaagctactccctccagcaat--tgtactaaatggagactgagatgct 417		
Dd	596 GTGATGACAGAATCTCTTCTCATCCACCATCGGTGACCTGGAATGGGATCTGGTAGT 655		
OY	418 gattatcagtcagtaaacatcagtggttcaatcttcbactctgcatgtagatcgtggtgga 477		
Dd	656 GACCTCATATCAGTACTTTCAGTGGCTAAATTTGTATCATGCGTGGAAATGATGTTGGGA 715		
OY	478 ggcatacagtggtgcacatgctcagaaaggt 508		
Dd	716 GGCACTTTAGSGCGTTTCATTATCAAGACAGGT 746		
RESULT 11			
RNDUIR	2597 bp mRNA linear ROD 30-MAR-1998		
LOCUS	Rattus norvegicus mRNA for putative integral membrane transport protein (UST1r).		
DEFINITION	protein (USTR1r).		
ACCESSION	V09945		
VERSION	V09945.1 GI:3004481		
KEYWORDS	Integral membrane transport protein; UST1r gene.		
SOURCE	Norway rat.		
ORGANISM	Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
REFERENCE	1 (bases 1 to 2597) Schomayr,E., Spitzzenberger,F., Engelhardt,M., Martel,F., Ordig,N. and Grundemann,D. Molecular cloning and characterization of two novel transport proteins from rat kidney FEBS Lett. 425 (1), 79-86 (1998)		
TITLE	JOURNAL MEDLINE 2 (bases 1 to 2597) Grundemann,D. Direct Submission Submitted (10-DEC-1996) D. Grundemann, University of Heidelberg,		

FEATURES		Location/Qualifiers	
source	gene	cds	
Heidelberg, FRG			
			1..2597
			/organism="Rattus norvegicus"
			/db_xref="taxon:10116"
			/tissue_type="kidney"
			87..1745
			/gene="UST1r"
			87..1745
			/gene="UST1r"
			/codon_start=1
			/product="putative integral membrane transport protein"
			/protein_id="CAA71076.1"
			/db_xref="GI:3004482"
			/db_xref="SprEmbl:070609"
			/translation="MAFDLLNVOYSGLSGRFOILLMTPLILNIIISPSLLENFTAVY FNHRKWPVILDNDFVSGNDNLSIQDILLRVSILPDSLRPREKRRVQPOWDLILN GTSSVSEPEDEPCVDGWNVDQSTFLSYIIEMDVSOSLISIAKFLFGLIVG ILVPLDRFGRRLILICASLOMAVTECAFAFPLFLYCSLRLAGISFETLNS LLILEMTPKFOALATGLICAGAIQGVLAGLAFVYNNHHLNLSVPIFFLVP RMLESARMLINTKLGKLEIKVAHINNSKSTDLVTEVVRTIMELELSOR SLMDLPEPTPLRRKICLSLVRVYVSVGLINQHLINWFLIQCLGITTTP NLVGRFLVNHGRRISQFLTSLFGISLAIIFPEQKQILRMVLATFGVFSVS SALVHNDLLEPTILIRATLGVIGIAGSGLALSPFLMLIRYSLSLWIIYGLSFLG GLVLLLEPTNQLPDSIDQVENEGRASRQGRKNDLIRKVTQF"
BASE COUNT	743 a	550 c	528 g 776 t
ORIGIN			
Query Match	37.3%	Score 300:	DB 10; Length 2597;
Best Local Similarity	74.4%	Pred.	No. 5.6e-76;
Matches 378;	Conservative 0;	Mismatches 130;	Indels 0; Gaps 0;
QY	1	atgagccttgagagccttgagcaagttgagagccttgagagattcagatgtcat	60
DB	87	ATGGCCTTCAGGACCTCTGATTAATTTGGACGCTGGGAGGTTCCAGATCCTTAG	146
QY	61	ctggttttattctctccctctctatgttattaaaccctatatactgtagaacttt	120
DB	147	ATGACTTTTCATTTGATCTTCAACATATATATTTCTCTCATTTCTATGGAGAACTTC	206
QY	121	gctgcagacccttcgctgcacatcgttgcgtggttccacaatgctggagacaataactgatact	180
DB	207	ACAGCAGCATTTCCCAATCATATGCTGCTGGGTCCCCATCTTCGAAATGACACTGTCCT	266
QY	181	ggtaatgaacttggaaatcctcagtgaagatgcccctcttgagatcctatccacaagac	240
DB	267	GGTATATGACATGAGGAACTGACGCAAGATGACCTCGTGAAGGAGCTCATGCCCTGGAC	326
QY	241	tcaaatcggagccaaggaagtgcgtctgctttgttccaatcccacagtggcagcttcctaac	300
DB	327	TCCGATCTGAGAACCAAGAAAGTGTGCTGCTTTTCCAAACACAGAGTGGATCTCTTCAT	386
QY	301	ctgaatggagactatccacagacaaagttaggacagacacagacccttgtagatgctcg	360
DB	387	CTAAGCGGTACTTTTCCAGTGTGACAGAGCCACACACAGACCCCTGTGTGATGCTGG	446
QY	361	gtatatgataaagctactctcccttgcacaattgtgaactaagtggagacctgtatgtgat	420
DB	447	GTTGATGACCAAAACACACTTCTTCTCCACCATTTATACCGAGTGGACCTGGTGTGAA	506
QY	421	tatagtcacactgaataatgaagtggttccaattcctacttctgagatgaatgctggtggagcc	480
DB	507	TCTAGTACACTGGGATTAATATGCTTAATTTCTATTCTTGACTGGTATCTTAGTAGGAAT	566
QY	481	atcataagttggccatgtctcagacaggt	508
DB	567	ATCCTATATGGCCCTTAACAGACAGGT	594

LOCUS	AB051836	1995 bp	mRNA	linear	ROD 01-DEC-2000
DEFINITION	Rattus norvegicus mRNA for organic anion transporter 5, complete cds.				
ACCESSION	AB051836				
VERSION	AB051836.1	GI:17221619			
KEYWORDS					
SOURCE	Rattus norvegicus (strain:Sprague-Dawley) cDNA to mRNA, clone.lib:rnt.kidney.cDNA library clone:roAT5.				
ORGANISM	Rattus norvegicus; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Rodentia; Sciurognathi; Muridae; Murinae; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
REFERENCE	1 (sites)				
AUTHORS	Cha,S., Sekine,T., Ono,T., Hasegawa,M., Kanai,Y. and Endou,H.				
TITLE	Molecular cloning and characterization of multispecific organic anion transporter 5 expressed in the kidney unpublished				
JOURNAL	2 (bases 1 to 1995)				
REFERENCE	Cha,S.				
AUTHORS	Direct Submission				
TITLE	Submitted (28-NOV-2000) Seok Ho Cha, Kyorin University School of Medicine, Department of Pharmacology and Toxicology; 6-20-2 Shinkawa, Mitake, Tokyo 181-8611, Japan				
JOURNAL	(E-mail:nal90@kyorin-u.ac.jp, Tel:+81 422 47 5511(ex.3453), Fax:+81 422 79 1321)				
FEATURES	Location/Qualifiers				
source	1..1995				
	/organism="Rattus norvegicus"				
	/strain="Sprague-Dawley"				
	/db_xref="taxon:10116"				
	/clone="roAT5"				
	/clone_lib="rat.kidney.cDNA.library"				
CDS	121..1776				
	/codon_start=1				
	/product="organic anion transporter 5"				
	/protein_id="BAB78471.1"				
	/db_xref="gi:17221620"				
	/translation="MARQDLITQIGCGRFQILHLIPIYLIQFIIVPPTVELENTAA /translition-"MARQDLITQIGCGRFQILHLIPIYLIQFIIVPPTVELENTAA PSHRWPEILDNNTSDNNRSRLISODDLIRLSPDSSLRPREKCRRTIQOPMDLHNL GTFSTVIEDPEPCVDGWNVDPQSFTLSLTVOIMDVCSOTLNSVAKTIFMGJLIG LMGHLDKRFGRFKLYTCGLTLAVETESVAFAPELLCSIRFLTGSSICIRINS LLILEMTSPKFOAMMALIFSGAIGOVLLVAFGLINMOHLQIAMSVPVFLLIP RLMSARWMLITTNKPQEGIKELIVAINIKNSROYLTEVVYTKKDELAETP PSPLVTRTPNLRKRICCLISFYRCVSLISTGILLINOYLSNKVFLCLOGVCTPPL NLGNFSNMNGRRRTQIIEMSVAGISLISLSTFFIOEQOIRIVIASGGALSSALL STVALSNELVPIVIRATAVGIVGIFGSAAISPULMLMTFSLSLPWITYGVLPILS SLVVLIDEPENROPIDPSIDIOVENRKSRSEVKKDVAKVTF"				
BASE COUNT	559 a 448 c 412 g 576 t				
ORIGIN					
Query Match	36.5%; Score 293.6; DB 10; Length 1995;				
Best Local Similarity	73.6%; Pred.No.4.1e-74;				
Matches 374; Conservative	0; Mismatches 134; Indels 0; Gaps 0				
Oy	1 atggcccttgaggaagccttggatcaagtgtgaggccttgaggagatttcagatgctcat 60 Db 121 atggcctttcaggaccttttaactcaaatgttgctgcctggcgagagttccagatcctttac 180 Oy 61 ctggttttatcttccttcctctcctcctcatgltatlaaacctccatactactgtagagaactt 120 Db 181 ttgatTTTTGTACTGATGCATTGCTTTATCCCGTAGTCCACACTGTTTGGGAACCTTC 240 Oy 121 gctgcagccattccctggtgatogtgtgctggtgtccaaatgctggaacaataataatgatat 180 Db 241 actgcagccattccctaagtcacgcttttgagggtcccattccctgcacataaaccattatgctt 300 Oy 181 agtaagaacvtgaatcctcagctgaagatgccctctttagaattcctcltccccactagac 240 Db 301 gatTAATAAACAGTAGATCCTCAAGCCAAGAACCTCTTAAGATGCCATCCCAATGGAC 360 Oy 241 tcaaatctgaagccagagaagtgtcgctgcttgtccatccccaagtcgagcagctcttacc 300 				

QY	361	TCCAACCTGAGGCGCAGANAATGCGGTGCTATATCCAAACACAGAGTGGAGTCTCTTCAT	420
QY	301	ctgaatggactatccacagcaacaagtgaaggcagacacagaacctgtgtgtagtgcctg	360
Db	421	CTGATATGGACACCTTCTCCACGTATACAAAGACGACACAGAGCCCTGTGTGATGGTGG	480
QY	361	gtatattgataaagtacttccctcttcgacacattgtgactaagtgtggaacctgtgtgat	420
Db	481	GTCATATGCCAAAGACACTTTCCTCCACACAGAGACCCAGTGGGACCTGTGTGTGGA	540
QY	421	tatcgaatcctgaatcagtggttcaattcctcactcttcgaactgtgaagtcggtggagac	480
Db	541	TCTCAAAACACTGAATTCAGTGGCTAAATTATCTACATACACTGGATCTTATTAAGGCAT	600
QY	481	atcctagtggtggccatgctctcagaccggt	508
Db	601	CTCATGGGTGGCCATTATTCAGACAGAT	628
RESULT 13			
AK057654			
LOCUS	2253 bp	linear	PRI 31-OCT-2001
DEFINITION	Homo sapiens CDNA FLJ31092.f1s, clone TRACH2000665, moderately similar to Rictus norvegicus mRNA for putative integral membrane transport protein (USTIR).		
ACCESSION	AK057654		
VERSION	AK057654.1	GI:16553416	
KEYWORDS	oligo capping; f1s (full insert sequence).		
SOURCE	Homo sapiens trachea CDNA to mRNA, clone_lib:"TRACH2 clone:"TRACH2000665.		
ORGANISM	Homo sapiens		
REFERENCE	Emmavrota, Metaxoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (sites)		
	Suzuki,O., Sasaki,N., Aotsuka,S., Shoji,T., Ichihara,T., Shiohata,N., Matsumoto,K., Hirano,M., Sano,S., Nomura,R., Yoshikawa,Y., Matsumura,Y., Moriya,S., Chiba,E., Momiyama,H., Onogawa,S., Kaeriyama,S., Satoh,N., Matsunawa,H., Takahashi,E., Kataoka,R., Kuga,N., Kuroda,A., Satoh,I., Kamata,K., Takami,S., Terashima,Y., Watanabe,M., Sugiyama,I., Irie,R., Otsuki,T., Sato,H., Ota,T., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Matsushima,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahata,K., Masuko,Y., Nagai,K. and Isogai,T.		
TITLE	NEO human cDNA sequencing project		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2253)		
AUTHORS	Isogai,T., Otsuki,T. and Sugiyama,T.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1552-3 Yana Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp) Tel:81-438-52-3951, Fax:81-438-52-3952		
COMMENT	NEO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.		
FEATURES			
source	Location/Qualifiers		
	1..2253		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="TRACH2000665"		
	/tissue_type="trachea"		
	/clone_lib="TRACH2"		
	/note="Cloning vector: PME18SF13"		
	208..1866		
	/note="unnamed protein product"		
	/codon_start=1		

/protein_id="BAB71543.1"
 /db_xref="GI:16553417"
 /translation="MAFQELLNQGSGRFOILQISFVLFVLGVLPVPHITMENTAAI
 PNHRMABILDNDASDGRSRLSODDLIRISIPDSNRILDKCRFPQPMHLLH
 GTFSENEPEPCVDGWYDRSNELSTIVEMDVCESOLNASKSPSPMGLFGA
 IICGHLSDRFKFLITCALOPATTECAFPAPFSEYICLLRPLAGSVEPISNSH
 LLMLEWSPKFLGVAVTSCASISGMIIAGLAFLEFIMRLQIAMSKEPFLIT
 RMMSARMLIVTKPKQGLKELKRVAMHNMKNSGNLTMEVEMASKELEAKRK
 SSPDLFHTPLIRKICVLSFMRLEFISGLHLOHSTNIILOFSLGIL
 SVIGHFVLMHGRRIQVLMISLRIGEMLTAVFVPOEMOTIRIIMATLAALSICM
 VSNHINELLPTTIRATGMVIAFMGNSGLFLAPFLMLATVYSANLPIFYGFSILY
 GFSEVLLPEFTKNOPLPDSIHVDGNMKSROAKEDSDIIVKTRF"

BASE COUNT 629 a 460 c 455 g 709 t

ORIGIN

Query Match 36.5%; Score 293.6; DB 9; Length 2253;
 Best Local Similarity 73.6%; Pred. No. 4.1e-74;
 Matches 374; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

OY 1 atggccttggaagccttgtagtaagttggagccttggaagattcagatgctcat 60
 DB 208 ATGGCCTTTCAGGAAGCTCTGATCAAGTTGGAGCCTAGGAAGATTCAGATCCTTCAG 267
 OY 61 ctggttttattctccctctcctcatgttattatccctcatatactgctagagaatt 120
 DB 268 ATCCTCTTGTCTCTTCTCTGCTAGCCTTGGTACTCATATTACATGAGAACCTTT 327
 OY 121 gctgagccattctctgctcatcgtgctgggtccacatgctggacaataactgact 180
 DB 328 ACTCGAGCATTTCCCAATCATCTGCTGGGCCCATCTTGACCAATGACACCTGCCT 387
 OY 181 gtaataagactggaactcctcagtaagaatgccctctgagaattctatccactagac 240
 DB 388 GATATGCGCACTAGGATCTTGAGCCAAAGATGACCTCGAGGATTCATCCCTCGAT 447
 OY 241 tcaaatctgaagccagaaagtgtcgtcgtctgcttccatcccaatgagcctcttcac 300
 DB 448 TCCAACTGAGACAGTGGATAAATGTCGTGTTTGCCCAACCAATGGCATTTCTCAT 507
 OY 301 ctgaatggactatccacagacaaagtgaagcagacagaaacctgtgtgagatgctgg 360
 DB 508 TTGATATGCGCACTTCTCCAAAGACAGACAGCCAGACCTGATGATGATGCTGG 567
 OY 361 gtaataagactggaactcctcctcagccttgtagtaagttggagccttgtagat 420
 DB 568 GTGATGACAGAGCACTCTCTTTCATCAATGAGCTGAGTGGACCTGTGTGAA 627
 OY 421 tataatcactgaatcagtggttcaatctcctactctgactggaatcgtgtggaagc 480
 DB 628 TCTCAGGACGATTTGTTGCTAAATTTTCATTCATGATGGCTTTTATAGGGCT 687
 OY 481 atcataagtgagcctatgctcagacagct 508
 DB 688 ATCATATGTGGCCATTTGTCAGACAGGT 715

RESULT 14

BC021384 1964 bp mRNA linear ROD 18-JAN-2002
 LOCUS BC021384
 DEFINITION Mus musculus, clone MGC:29260 IMAGE:5064542, mRNA, complete cds.
 VERSION BC021384.1 GI:18204002
 KEYWORDS MGC.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1964)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (14-JAN-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer

REMARK

COMMENT

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
 Sequencing Center (NISC),
 Gaithersburg, Maryland,
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc-mgc@nrl.nih.gov
 Shenoy, K., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
 Benjamin, B., Blakesley, R.M., Bouffard, G.G., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-T., Karlins, E., Legaspi, R.,
 Lim, M., Maduro, O.L., Mastello, C., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Snyder, B., Stancir, S., Thomas, P.J.,
 Tjongson, E.E., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A.,
 Zhang, L.-H., and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAC Plate: 39 Row: n Column: 1
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis, similarity but not identity to protein.

FEATURES

source

1. 1964
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="MGC:29260 IMAGE:5064542"
 /tissue="type="Kidney, normal. 5 month old male mouse."
 /clone_lib="NCI CGAP_Kid4"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 84. 1739
 /codon_start=1
 /product="unknown (protein for MGC:29260)"
 /protein_id="AAH21384.1"
 /db_xref="GI:18204003"

CDS

/translation="MAFQELLNQGSGRFOILHMFVILCHALSAPTILMENTAAI
 PSNRHABILDNDASDGRSRLSODDLIRISIPDSNRILDKCRFPQPMHLLH
 GTFSENEPEPCVDGWYDQSTFLSTTVIOMDVCOSLNSVAKTIYMTGLTGY
 IMGHLSDFKFKLIFTCSLDKMAITTEVCAFAPSEFLRSLRSLGFSSTLRNSA
 LLIEMWSPKFOALVMAIFIASVGQFLIGLAFAPFNHMLDLALSPMFLILPT
 RMLESARMLIMARKPORSLEKKAACVNRKNSGDALTLEVAVTKIKELAEAOFK
 PSPDLFETPMRKRIKLSFVRVSVSLGILINIOYLSNNVFILOCIAGVACIPA
 NVLGNFSNNVGRMTOLIMSVYIGISITLAVPPEOMQILRVLSTGGAISASST
 STVHANELVYTIIRATRLGVYGIAGSAGSLSPMLMLTTYSASLFWITLYGILPFLG
 GLVALPEFTKNOPLPDSIODIENKRRSSKRAKRVYAKVPL"

BASE COUNT 536 a 456 c 406 g 566 t

ORIGIN

Query Match 36.1%; Score 290.4; DB 10; Length 1964;
 Best Local Similarity 73.2%; Pred. No. 3.4e-73;
 Matches 372; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

OY 1 atggccttggaagccttgtagtaagttggagccttggaagattcagatgctcat 60
 DB 84 ATGGCCTTTCAGGAAGCTCTGATCAAGTTGGAGCCTAGGAAGATTCAGATCCTTCAG 143
 OY 61 ctggttttattctccctctcctcatgttattatccctcatatactgctagagaatt 120
 DB 144 ATGATATTTGTGTAATCTGCATGCTGTCGCTCTCACACACTTTTGGAGAACTTC 203
 OY 121 gctgagccattctctgctcatcgtgtcgtgggtccacatgctggacaataactgact 180
 DB 204 ACTCGAGCATTTCTAGCATCTGCTGGGCCCATCTTGACCAATGACACCTGCCT 263
 OY 181 gtaataagactggaactcctcagtaagaatgccctctgagaattctatccactagac 240

Db 264 GATAATGCTAGTAGATCTGAGCAGCAAGTACCTCTGAGAGTCTCCATCCOCCCTGGAC 323
 QY 241 tcaaatctgaagccagagaagtgctgcttgctccatcccaagtggagcttcttcac 300
 Db 324 TCCAACTGAGACCGGATTAATGCGCTATATTCACACAGTGGATCTCTTCAT 383
 QY 301 ctgaatggactatcccaagcaagtgaggagacagaaacctgtgtgagtgctg 360
 Db 384 CTGAATGGACCTTCCCTCAGCTGTATACAGACACAGAGCCCTGTGTGGATGGCTGG 443
 QY 361 gataatgaatcaagctacttccctcgaacatgtgactaagtgaggacctgtatgat 420
 Db 444 GTGATATGACCAAGACACTCTCTCTCCACACAGTACCCAGTGGAGCTGTGTGGA 503
 QY 421 tacaatctgaagccagagaagtgctgcttgctccatcccaagtggagcttcttcac 480
 Db 504 TCTCAGGACCTGAATTCAGTGTGCTTAATTAATACATGATGATCTTTATAGATAT 563
 QY 481 atcataagtgccatgtctcagacagagt 508
 Db 564 ATCATGGGTGGCCACTTATCAGACAGT 591

RESULT 15
 RNO132859 1911 bp mRNA linear ROD 04-JAN-2002
 LOCUS Rattus norvegicus mRNA for putative integral membrane transport protein UST4r.
 DEFINITION
 Rattus norvegicus mRNA for putative integral membrane transport protein UST4r.
 ACCESSION AJ132859.1 GI:18076048
 VERSION AJ132859.1
 KEYWORDS integral membrane transport protein; UST4r gene.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE 1 (sites)
 AUTHORS Honold,C., Gruendemann,D. and Schoemig,E.
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1911)
 AUTHORS Gruendemann,D.
 TITLE Direct Submission
 JOURNAL Submitted (08-FEB-1999) Gruendemann D., Dept. of Pharmacology, University of Heidelberg, INF 366, 69120 Heidelberg, 69120, GERMANY
 FEATURES
 source
 1..1911
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 /tissue_type="liver"
 121..1779
 /gene="UST4r"
 121..1779
 /gene="UST4r"
 /codon_start=1
 /product="putative integral membrane transport protein"
 /protein_id="CAC79639.1"
 /db_xref="GI:18076049"

gene
 CDS
 BASE COUNT 535 a 395 c 397 g 584 t
 ORIGIN
 Query Match 35.9%; Score 288.8; DB 10; Length 1911;
 Best Local Similarity 73.0%; Pred. No. 1e-72;

Matches 371; Conservative 0; Mismatches 137; Indels 0; Gaps 0;
 QY 1 atggccttgaggagcctcttgatcaagtggagagccttggaattccaatgctcat 60
 Db 121 ATGGCCTTTGAGGAGCTCTGTAATCAGTATGAGAGGCTTAGGAGATCTCCTCAG 180
 QY 61 ctgttttcttcttccctcctcctcctgataatccctatactctgagaactt 120
 Db 181 ATGCTTTTGTGTGCTTCCACAGTGTATGTGTGACTCATATATATAGAGAACTT 240
 QY 121 gctgaagccatctcgtgcatcgtgtgctgggtccaatgctggaataatctgatt 180
 Db 241 ACTGACCCATTTCCAGCACTGCTGCTGGGTCTTCTATCTCGACAAATGCTCTGTCT 300
 QY 181 gtaatgaactgaatctcgaatgaagtgagagtgcccttgagaatctcactcaag 240
 Db 301 GATTAACGACATAGATTTCTAACCAGATTAATCTCTGAAGATCTCCATCCCTTGGAC 360
 QY 241 tcaaatctgaagccagagaagtgctgcttgctccatcccaagtggagcttcttcac 300
 Db 361 TCCCACTGAGACAGTAAAGTGTGCTGTTTGGCCAAACAGTGGCATCTTGTGAT 420
 QY 301 ctgaatggactatcccaagcaagtgaggagacacagaaacctgtgtgagtgctg 360
 Db 421 TTGAATGATTAATCTCTCCAGTATTAATTAATTAATTAATTAATTAATTAATTA 480
 QY 361 gataatgaatcaagctacttccctcgaacatgtgactaagtgaggacctgtatgat 420
 Db 481 GTATATGACAGAGAGCACTTTCATCTACACTGAGACTGAGACTGAGACTGAGACTG 540
 QY 421 tacaatctgaagccagagaagtgctgcttgctccatcccaagtggagcttcttcac 480
 Db 541 TCTCAGGACCTGAATTCAGTGTGCTTAATTAATTAATTAATTAATTAATTAATTA 600
 QY 481 atcataagtgccatgtctcagacagagt 508
 Db 601 ATGTAATGGCTTCTCTCAGACAGT 628

Search completed: July 1, 2002, 15:41:50
 Job time: 10924 sec

EMBL: 297028; CAB09724.1; -
 DR InterPro: IPR003662; sub-transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 FT TRANSMEMBRANE; Transport; Glycoprotein; Ion transport.
 FT TRANSMEM 148 168 POTENTIAL.
 FT TRANSMEM 175 195 POTENTIAL.
 FT TRANSMEM 206 226 POTENTIAL.
 FT TRANSMEM 261 281 POTENTIAL.
 FT TRANSMEM 409 429 POTENTIAL.
 FT TRANSMEM 496 516 POTENTIAL.
 FT CAROHD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHD 95 95 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 562 AA; 61992 MW; 7DC1A67F32801D2D CRC64;

Query Match 31.2%; Score 439; DB 13; Length 562;
 Best Local Similarity 27.4%; Pred. No. 5.4e-34;
 Matches 99; Conservative 54; Mismatches 84; Indels 124; Gaps 5;

QY 1 MAFELLISQVGLGRFQMLHLVFLIPSLMLIPHLLENFAAIPGHRGVHMLDNTGS 60
 DB 1 MPEELLISQVGLGRFQMLHLVFLIPSLMLIPHLLENFAAIPGHRGVHMLDNTGS 60
 QY 61 GNETGISLSDALRISIPDLSNLRPEKCRFVHPQWOLHLNGTISTEADT----- 113
 DB 55 LSQLALSLSESL-ITVPLDAGKRCORRYAAPHMLGKNGTSGSLDADTESMDAA 113
 QY 114 -EPYDGVNVDOSYFPTVTKMDLVCDYSLKSVQFLLTGMVGIIIGVSDR-- 169
 DB 114 LQESDGSVSYSTVSTYSLISEMHLVCDMSFQKQGTIYMGVLYGALLFGLSDRYGR 173
 QY 170 ----- 169
 DB 174 RILLISNLAVSGTCAAFSSSSSLFCVFRGGLALSGLNTFSLIVEMIPRIRTA 233
 QY 170 ----- 184
 DB 234 VGTTCYCTIGQLLVLLVAFIRDMRWLTLAVSLPFVFFLIAMWFHSSRWLALSRT 293
 QY 185 DEGLKARKVARTGINKAEETLNIEVVRSTMOEELDAQRTKTYCDLFRNDSMKR-IC 243
 DB 294 EHALNKLSVARFNGRHEAEKLIKMLHESKKEKMSCTQSSITLDFNPAMKRRLC 353
 QY 244 I 244
 DB 354 I 354

RESULT 13
 ID 061185 PRELIMINARY; PRT; 545 AA.
 AC 061185;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE KIDNEY-SPECIFIC TRANSPORT PROTEIN.
 GN SLC22A6.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BAJB/C;
 RX MEDLINE=97197822; PubMed=9045672;
 RA Lopez-Nieto C.E., You G., Bush K.T., Barros E.J., Beier D.R.,
 Nigam S.K.;
 RT "Molecular cloning and characterization of NKT, a gene product related
 to the organic cation transporter family that is almost exclusively
 expressed in the kidney";
 RT J. Biol. Chem. 272:6471-6478(1997).
 CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).

-I- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC EMBL: U52842; AAC53112.1; -
 DR MGD: MGI:892001; SLC22a6
 DR InterPro: IPR003662; sub-transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 KW TRANSMEMBRANE.
 SQ SEQUENCE 545 AA; 59983 MW; CD3B762E5984933A CRC64;

Query Match 31.2%; Score 438.5; DB 11; Length 545;
 Best Local Similarity 28.7%; Pred. No. 5.8e-34;
 Matches 102; Conservative 52; Mismatches 77; Indels 125; Gaps 6;

QY 1 MAFELLISQVGLGRFQMLHLVFLIPSLMLIPHLLENFAAIPGHRGVHMLDNTGS 60
 DB 1 MAFNDLLKQVGVGRFQMLHLVFLIPSLMLIPHLLENFAAIPGHRGVHMLDNTGS 54
 QY 61 GNETGISLSDALRISIPDLSNLRPEKCRFVHPQWOLHLNGTISTEADT----- 120
 DB 55 AN-----LSKDGGLFARLPDKQGRPESCLFPPH-----NCT-EANGTGTPECLDGM 103
 QY 121 VYDOSYFPTVTKMDLVCDYSLKSVQFLLTGMVGIIIGVSDR----- 169
 DB 104 VIDNSTFPTVTKMDLVCDYSLKSVQFLLTGMVGIIIGVSDR----- 169
 QY 170 ----- 171
 DB 164 LQAVSGTCAAVAPNTVYCFRLSLGMSLAINCMILNEMMIFHRAVYGLIGYV 223
 QY 172 ----- 192
 DB 224 YSLGQFLAGIYAVPWHMLQLAVSPFVAFIYSWPFIESARWYSSSGRLDLTFLALQ 283
 QY 193 KVARINGINKAEETLNIEVVRSTMOEELDAQRTKTYCDLFRNDSMKR-IC 248
 DB 284 RVARINGKQEBGAKLSIEVLTSLQKELTLNKGASAMLELHCPILRLCLSLML 339

RESULT 14
 ID 096TC1 PRELIMINARY; PRT; 542 AA.
 AC 096TC1;
 DT 01-DEC-2001 (Tremblrel. 19, Created)
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
 DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE ORGANIC ANION TRANSPORTER 3.
 GN HOAT3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cha S., Sekine T., Kanai Y., Endou H.;
 RT "Molecular cloning and characterization of human organic anion
 transporter 3";
 RT Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB042505; BAB47393.1; -
 SQ SEQUENCE 542 AA; 59857 MW; 5053B36439554CCC CRC64;

Query Match 27.8%; Score 390.5; DB 4; Length 542;
 Best Local Similarity 25.4%; Pred. No. 2.6e-29;
 Matches 91; Conservative 46; Mismatches 70; Indels 151; Gaps 5;

QY 1 MAFELLISQVGLGRFQMLHLVFLIPSLMLIPHLLENFAAIPGHRGVHMLDNTGS 50
 DB 1 MTFEELLISQVGLGRFQMLHLVFLIPSLMLIPHLLENFAAIPGHRGVHMLDNTGS 60
 QY 51 VHMIDNNTGSGNETGISLSDALRISIPDLSNLRPEKCRFVHPQWOLHLNGTISTE 110
 DB 61 V-----LPMGPNGKPERCLRFVHP-----NMSLPDQTQ 89

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC TISSUE=KIDNEY.
 RX MEDLINE=99103997; PubMed=9887087;
 RA Hosoyamada M., Sekine T., Kanai Y., Endou H.;
 RT "Molecular cloning and functional expression of a multispecific
 RT organic anion transporter from human kidney.";
 RL Am. J. Physiol. 276:F122-F128(1999).
 RN [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=98433806; PubMed=9762842;
 RA Reid G., Wolff N.A., Dautzenberg F.M., Burckhardt G.;
 RT "Cloning of a human renal P-aminohippurate transporter, hROAT1,";
 RL Kidney Blood Press. Res. 21:233-237(1998).
 RN [3]
 RN SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RX MEDLINE=99137667; PubMed=9950961;
 RA Lu R., Chan B.S., Schuster V.L.;
 RT "Cloning of the human kidney PAH transporter: narrow substrate
 RT specificity and regulation by protein kinase C.";
 RL Am. J. Physiol. 276:F295-F303(1999).
 RN [4]
 RN SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RX MEDLINE=99160894; PubMed=10049739;
 RA Race J.E., Grassl S.M., Williams W.J., Holtzman E.J.;
 RT "Molecular cloning and characterization of two novel human renal
 RT organic anion transporters (hOAT1 and hOAT3).";
 RL Biochem. Biophys. Res. Commun. 255:508-514(1999).
 RN [5]
 RN SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RX MEDLINE=99393620; PubMed=10462545;
 RA Cihlar T., Lin D.C., Fritchard J.B., Fuller M.D., Mendel D.B.,
 RA Sweet D.H.;
 RT "The antiviral nucleotide analogs cidofovir and adefovir are novel
 RT substrates for human and rat renal organic anion transporter 1.";
 RL Mol. Pharmacol. 56:570-580(1999).
 RN [6]
 RN SEQUENCE FROM N.A.
 RP Bahn A., Prawitt D., Reid G., Enklaar T., Wolff N.A., Hillemann A.,
 RA Godelhardt S., Butler D., Knabe B., Schulten H.J., Gunawan B.,
 RA Fuesel L., Zabel B., Burckhardt G.;
 RT "Genomic cloning and characterization of the human renal organic anion
 RT transporter gene (hOAT1).";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MEDIATES THE NA(+)-INDEPENDENT TRANSPORT OF ORGANIC
 CC ANIONS SUCH AS P-AMINOHIPPURATE AND ALPHA-KETOGLUTARATE.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. BASOLATERAL
 CC MEMBRANE (POTENTIAL).
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: ISOFORM hOAT1-1 (SHOWN HERE)
 CC AND ISOFORM hOAT1-2; ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN KIDNEY. ALSO DETECTED IN
 CC BRAIN AND AT LOW LEVELS, IN SKELETAL MUSCLE AND PLACENTA. NOT
 CC EXPRESSED IN HEART, LUNG, LIVER, PANCREAS, COLON, THYMUS, SMALL
 CC INTESTINE OR PERIPHERAL BLOOD LEUCOCYTES. IN THE KIDNEY, STRONGLY
 CC EXPRESSED IN PROXIMAL TUBULE CELLS.
 CC -1- PTM: GLYCOSYLATED.
 CC -1- SIMILARITY: BELONGS TO THE SLC22 FAMILY OF TRANSPORTERS.
 DR EMBL: AF057039; AAC70004.1; -;
 DR EMBL: AB009697; BAA75072.1; -;
 DR EMBL: AF104038; AAD10052.1; -;
 DR EMBL: AB009698; BAA75073.1; -;
 DR EMBL: AF097490; AAD19356.1; -;
 DR EMBL: AF124373; AAD55356.1; -;
 DR EMBL: AJ249369; CAB77184.1; -;
 DR InterPro: IPR003662; sub.transporter.
 DR Pfam: PF00083; sugar_tr.1;
 KM Transmembrane; Transport; Alternative splicing; Ion transport;
 Glycoprotein.

FT TRANSMEM 136 156 POTENTIAL.
 FT TRANSMEM 185 205 POTENTIAL.
 FT TRANSMEM 249 269 POTENTIAL.
 FT TRANSMEM 338 358 POTENTIAL.
 FT TRANSMEM 396 416 POTENTIAL.
 FT TRANSMEM 426 446 POTENTIAL.
 FT TRANSMEM 485 505 POTENTIAL.
 FT CARBOHYD 39 39 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 56 56 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 92 92 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 113 113 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPPLIC 523 535 MISSING (IN ISOFORM OAT1.2).
 FT CONFLICT 14 14 G->S (IN REF. 3).
 FT CONFLICT 563 563 L->F (IN REF. 2).
 SQ SEQUENCE 563 AA; 61816 MW; 74AD3EA2678032E4 CRC64;
 Query Match 32.6%; Score 458.5; DB 4; Length 563;
 Best Local Similarity 29.2%; Pred. No. 7e-36;
 Matches 104; Conservative 51; Mismatches 82; Indels 119; Gaps 5;
 QY 1 MAPEELISQVGGIGRROMLHVFIPLSLMLIPHIILENFAAIPGRCHWMMDNNTGS 60
 DB 1 MAPNDLQOYGVGRQOQIOVTLVPLMLASHNTLNTAIPTHC-----RRPAD 54
 QY 61 GNETGISDALIRISIPLDNLNRPKCRFVHPWOLHNLNGTSTSEADTEPCVDGW 120
 DB .55 AN-----LSKNQGLEVLPLPRDQGPESCLRFSPQWMLPFLNGT-EANGTGATPCPDGW 109
 QY 121 VYDOSFPSTIYTKWPLVCDYQSLKSVYQFLTLGMLVGGIIGHVSDR----- 169
 DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 169
 QY 110 IYDNTFPTIYEMDVLGSHRALROLASLVYGVLLGAMVGYLDRGRKKVLLINY 169
 QY 170 -----WL----- 171
 DB 170 LQTVASGTAAPNPPIYCAFRLLSGMALAGISLNCMTLNVEMPHITACVGTLLGY 229
 QY 172 -----VESARWLLITNKLDEGLRALR 192
 DB 230 YSLGQFLGVAVAPVPHRHQLQLLSAPFAFYISWFFESARWSSSGRLDTLRALQ 289
 QY 193 KVARNGIKNAEETLVIEVRSIMOEELDAQTKTYCDLFRPMSMKRRCIIVFL 248
 DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 248
 DB 290 RVARINKREKAKLSMEVLRASLOKELTMKGQASAMELLROPTLRHLFLCSML 345
 RESULT 10
 ID 035956 PRELIMINARY; PRT; 551 AA.
 AC 035956;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)
 DE 01-JUN-2001 (TREMBLrel. 17, last annotation update)
 DE RENAL ORGANIC ANION TRANSPORT PROTEIN 1.
 GN SLC22A6 OR ROAT1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RN SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
 RC TISSUE=KIDNEY;
 RX MEDLINE=98043701; PubMed=9374486;
 RA Sweet D.H., Wolff N.A., Pritchard J.B.;
 RT "Expression cloning and characterization of ROAT1. The basolateral
 RT organic anion transporter. In rat kidney.";
 RL J. Biol. Chem. 272:30088-30095(1997).
 RN [2]
 RN SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
 RP STRAIN=SPRAGUE-DAWLEY; TISSUE=KIDNEY;
 RX MEDLINE=97373539; PubMed=9228014;
 RA Sekine T., Watanabe N., Hosoyamada M., Kanai Y., Endou H.;


```

OY 121 YDQSYEPSTIYTKWDYICDQDQSLKSVYQFLLTGLMVGIIIGHVSDR 168
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 YDHSYFSTIYTTWDLVCSOALRPMQOSIFLGIILGAVCCHASDRGRVRLTWSY 188
    ----- 168
OY 170 ----- 168
Db 181 LLSVSGTAAAFMPTFLPYCLFRLLASAVAGVMNNTASLMEWTSAQSGDLYVTLNALG 240
    170 -----
OY 170 -----
Db 241 FSPGQVLLGSAVYGRSRHMLQLAVSAPELFEFYVSMLEPSASRWLTIVYKLDQIGLQ 300
    -----
OY 193 KVARTNGIKNAEETLNTIEVASTQOEELDAQOTKTVCDFRPNDSMKRRICI 244
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 RVAAVNRRAKAGDTLTMEVLRSAAMEEPPSRDKAGASIGTLTLHPFGRLHRKRII 352
    -----

```

```

RESULT 7
Q9NOC2 PRELIMINARY; PRT; 506 AA.
ID Q9NOC2:
DT 01-OCT-2000 (TRMBLrel. 15, Created)
DT 01-OCT-2000 (TRMBLrel. 15, Last sequence update)
DE 01-DEC-2001 (TRMBLrel. 19, Last annotation update)
DE PUTATIVE ORGANIC ANION TRANSPORTER.
OS OAT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Carnivora; Homnidae; Homo.
RN NCBI_TaxID=9606;
RP [1]
RC SEQUENCE FROM N.A.
RF TISSUE=KIDNEY;
RA MEDLINE=20422308; PubMed=10964714;
RA Bahn A., Pirwitz D., Buttler D., Reid G., Enklaar T., Wolff N.A.,
RA Edlinghaus C., Hillemann A., Schulten H.J., Gunawan B., Fuenesi L.,
RA Zabel B., Burchardt G.
RT "Genomic structure and In Vivo expression of the human organic anion
RT transporter 1 (hOAT1) gene."
RL Biochem. Biophys. Res. Commun. 275:623-630(2000).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY) .
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL: A2251529; CAB94830.1.
DR InterPro: IPR003662; sub-transporter.
DR Pfam: PF00083; sugar_tr_1.
DR Transmembrane.
SQ SEQUENCE 506 AA; 55858 MW; D8EBAE8A113E6C5E CRC64;

```

Query Match	32.6%	Score 458.5	DB 4	Length 506
Best Local Similarity	29.2%	Pred. No. 6.1e-36		
Matches 104;	Conservative 51;	Mismatches 82;	Indels 119;	Gaps 5

[illegible]

QY 193 KVA^{NT}NGIKNAE^TLIN^EEV^RST^{MO}EE^{LD}AA^{OT}K^{TT}Y^{CD}LFRNDSM^{KR}IC^{IL}V^{FL} 248
 :||| | :||| :||| :||| : : : :||| |
 Db 290 RV^{AR}IN^KK^{RE}EG^{AK}LME^{VL}RA^{SL}LN^{KE}LL^{MG}QASMAE^{LL}RC^{PT}L^{HE}FL^{CL}SM^L 345

RESULT	8	
ID	09NOA6	
AC	09NOA6	PRELIMINARY;
DT	01-OCT-2000 (TREMblrel. 15, Created)	PRT; 519 AA.
DT	01-OCT-2000 (TREMblrel. 15, last sequence update)	
DT	01-DEC-2001 (TREMblrel. 19, last annotation update)	
DE	PUTATIVE ORGANIC ANION TRANSPORTER.	
GN	homo-4.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=KIDNEY;	
RX	MEDLINE=20422308; PubMed=10964714;	
RA	Bahn A., Pirwitz D., Buttler D., Reid G., Enkelaar T., Wolff N.A.,	
RA	Edinghaus C., Hillemann A., Schulten H.J., Gunawan B., Fuenesi L.,	
RA	zabel B., Burchardt G.;	
RT	"Genomic structure and in vivo expression of the human organic anion	
RT	transporter 1 (hOAT1) gene.";	
RL	Biochem. Biophys. Res. Commun. 275:623-630(2000).	
DR	EMBL; AJ271205; CAB97249.1; -	
SQ	SEQUENCE 519 AA; 57357 MW; EL748CF9E2002P2 CRC64;	

Query Match	32.6%;	Score 458.5;	DB 4;	Length 519;
Best Local Similarity	29.2%;	Pred. No. 6.3e-36;		
Matches 104;	Conservative 51;	Mismatches 82;	Indels 119;	Gaps 5

QY	1	MAFEELLSOVAGISGEPOMLHLVETILPRLSMTLLPHLLENFNAALIPGRHCWMMDNNTGS	60
Db	1	MAFNLDLQOVGCVGRFOQIOYTLVLYPLTLMASSHNTLFNAALPTHTHC-----RPPAD	54
QY	61	GNFGTILSEDLALRISIPLDNSLRPERKCRREVPWQMDLHNGTISHSTSEADTEPCYGM	124
Db	55	AN-----ISKNGLELWVPRDROQPESECLRFTSPQWGLPFLNGT--EANGGATEPCIDGM	100
QY	121	VYDOSYEPSTIVTKMDLYCDYOSKSVQFOLLGMYGGIIGSHVDR-----	166
Db	110	IYDSTPEPSTIVTEMDLYCSHRALRQLAQSILYMGVILLGMAVFESTLADLRGRKRYLTLY	166
QY	170	-----WL-----	171
Db	170	LQNAVSGTCAAFAPNPFITYCAFLLSGMALAGISLNCMTLNVEMPIHTRACVOTLIGYV	229
QY	172	-----*ESARWLITITKRLDEGKALR	192
Db	230	YSLGQELLAGVAYAVPMRHLLQLVSAPEFAFEIYSWFFETESARWSSSSGRDLTLRALQ	289

RESULT 9
095742 PRELIMINARY; PRT; 563 AA.
AC 095742; 095187; 09UE08; 09UEG5;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE REVAL ORGANIC ANION TRANSPORT PROTEIN 1.
GN SLC22A6 OR OAT1 OR ROAT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

DR EMBL: AC044/90; AAN06133.1; 233BE6CA520E58A CRC64;
SQ SEQUENCE 550 AA; 59971 MW;

[illegible]

```

QY 61 GNETIILEEDLRLRSTPLDSENLRPEKRCRRVHOOMLHLNGLNTHSTSEADTEPCYDGM 120
Db 61 GNDNNGNSODDLLRVSTPLDSDLRPEKRRVVOOMLHLNGLNTHSTSEADTEPCYDGM 120
QY 121 VYDOSYFPTSTVTKMDLVCDYQSLKSVYQVLELLTGMVLGGIIGHVSD----- 168
Db 121 VYDOSYFPTSTITTEMDLVCEOSQSDISIAKEFLTGLIVGNILVGPDLDFGRRLILICAS 180
QY 169 ----- 168
Db 181 LQMAVTECAAFAPFTFLIYSLRFLAGISFTYVLNSALLIEMTRKPFQALATGILLCA 240
QY 169 ----- FVLVESARWMLITNKLEDEGLKAR 192
Db 241 GAIQGVLAGIAFTVRNWHHLHLAMSVPIFELLVPTIRMLSESARWMLIMTKLQKGLKEI 300
QY 193 KVAQTNGIKNAEETLIEVYRSTMOEFLDAOAKRTTYCCDFRAPSMBKRRTICLIVFLR 249
Db 301 KVAHINQMKNSDVLTEIVERTIMKEELSESQTKSSIMDLFERPDLNKRKRCLISLVR 357

```

RESULT	2	
096LK3		
ID	096LK3	PRELIMINARY; PRT; 552 AA.
AC	096LK3;	
DT	01-DEC-2001 (TrEMBLrel. 19, Created)	
DT	01-DEC-2001 (TrEMBLrel. 19, last annotation update)	
DT	01-DEC-2001 (TrEMBLrel. 19, last annotation update)	
DE	CDNA FLJ33092. f15. clone TRACH200065, MODERATELY SIMILAR TO RAT TUS	
OS	NOVORDECS PUTATIVE INTEGRAL MEMBRANE TRANSPORT PROTEIN (US1TR).	
OC	Homo sapiens (human).	
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
CC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxId:9606;	
SN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=TRACHEA;	
RA	Suzuki O., Sasaki N., Aotsuka S., Shoji T., Ichihara T., Shiohata N.,	
RA	Matsumoto K., Hirano M., Sano S., Nomura R., Yoshikawa Y.,	
RA	Kaeriyama S., Satoh N., Matsunawa H., Takahashi E., Kataoka R.,	
RA	Kuga N., Kuroda A., Satoh I., Kamata K., Takami S., Terashima Y.,	
RA	Watanabe M., Sugiyama T., Irie R., Otsuki T., Sato H., Ota T.,	
RA	Makamatsu A., Ishii S., Yamamoto J., Isono Y., Kawahito Y., Saito K.,	
RA	Nishikawa T., Kimura K., Yamashita H., Matsuo K., Nakamura Y.,	
RA	Sekine M., Kikuchi H., Kanda K., Wagatsuna M., Murakawa K.,	
RA	Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,	
RA	Suzuki Y., Sugeno S., Nagahari K., Masuno Y., Nagai K., Isogai T.,	
RT	"NBD human cDNA sequencing project."	
RL	Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.	
SR	EMBL: AK075654; BAB71543.1.	
SQ	SEQUENCE 552 AA; 62140 MW; 39A831C0FD26CA68 CRC64;	

Query Match	55.0%	Score 773	DB 4	Length 552
Best Local Similarity	43.4%	Pred. No. 2.2e-66		
Matches 155	Conservative 41	Mismatches 53	Indels 108	Gaps 1

[illegible]

```
OY    169 -----MWIVESARWLLITTKDEGKALR   122
Db     241 ASIGTMILAGLAFPRIRRHQLAMSVPIPEELLTRMSESARWLTVTKPKQKGLELR   300
OY    193 KVARNGIKNAEELNTEVVSZMOEDLAQOTYTGCDFRNPSMKRRICILEFLR   249
      OY    |||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db     301 KVAHMGKKSGNTLTMEVVASAKNKLEAAKRSSRPDLPHTLTCKRKICVLJSFMR   357
```

RESULT	3		
091WJ2			
ID	091WJ2	PRELIMINARY;	PRT; 552 AA.
AC	091WJ2;		
DT	01-DEC-2001 (TREMBLrel. 19, Created)		
DT	01-DEC-2001 (TREMBLrel. 19, last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, last annotation update)		
DE	SIMILAR TO SOLUTE CARRIER FAMILY 22 (ORGANIC CATION TRANSPORTER)-LIKE		
DE	2.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxId=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISD6-KIDNEY;		
RC	TISD6-KIDNEY;		
RA	Strausberg R.;		
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC014805; AAH14805.1; -.		
SQ	SEQUENCE 552 AA; 61360 MW; F5D2498FC673E8FC CR664;		

Query Match	53.4%	Score 751	DB 11	Length 552
Best Local Similarity	41.5%	Pred. No.	3e-64	
Matches 148	Conservative 44	Mismatches 57	Indels 108	Gaps 1

QY	1	MAFEELISVGSGGRPMOLVLPISIMILPHILLENPAALPCHRCVHMIDNNGS	60
QY	1	MAFEELISVGSGGRPMOLVLPISIMILPHILLENPAALPCHRCVHMIDNNGS	60
Db	1	MAFDOLLNLVSGHGRQIOLMAFLLLCNVITMHTVLENFTANIGHRCVHIIDNDNS	60
QY	61	GNETGISSEDALLRISIPLDNSLRPERKRRFHPWOMLLHNGTITHSTSEADTEPCVDGM	120
Db	61	NNNSGILSODDLRLRISIPLDNSLRPERKRRFHPWOMLLHNGTITHSTSEADTEPCVDGM	120
QY	121	VYQSGSPSPITYKMPVCDYQSLKSVQOELLTGLVIGIIGGHVSD-----	168
Db	121	VYQRFPLSTTYEMNLVCGSELNSVAFIFLGLVLAGHFVGHLSDKFGKRLFRCAL	180
QY	169	-----	168
Db	181	LQWATGTCTALAPRFIFYCLRLFTGLCIPIPNNSVYLMLEMTSPKQALVYTLSSMS	240
QY	169	-----RWVESARWLITTKIKIDEGKALR	192
Db	241	HNFGGILLAGLAPRPNMHNHLDAISVPIVLLIPLTFRMLTESARWLITTKRFOKALDELR	300
QY	193	KVARTNGINKAEELINIEVVRSTQMEELDAQTKTTCVCLFENPSPRRKRICILVPR	249
Db	301	KVASKNGIKNSDVLTMELVYVRTIMDELAIPRTKSLMDLFIEMPRKRRICILCLR	357
RESULT	4		
Q96S37			
ID	Q96S37	PRELIMINARY;	PRT, 553 AA.
AC	Q96S37;		
DR	01-DEC-2001	(TREMBLrel. 19, Created)	
DR	01-DEC-2001	(TREMBLrel. 19, last sequence update)	
DR	01-DEC-2001	(TREMBLrel. 19, last annotation update)	
DE	RST		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
NC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.		
NCBI_TaxID	9606;		

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 1, 2002, 11:39:10 ; Search time 26.48 Seconds
(without alignments)
1750.855 Million cell updates/sec

Title: US-09-674-235-1

Perfect score: 1406
Sequence: 1 MAFFELLISVGLGRFQMLH.....RKISRKRHKNDCTKTKTF 268

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 17294929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: SPREMBL-19:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mammal:*
8: sp_mhc:*
9: sp_organelle:*
10: sp_phage:*
11: sp_plant:*
12: sp_rodent:*
13: sp_virus:*
14: sp_vertebrate:*
15: sp_unclassified:*
16: sp_virus:*
17: sp_bacteriophage:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	818	58.2	552	11 070609	070609 ratu
2	773	55.0	552	4 0961X3	0961X3 homo sapien
3	751	53.4	552	11 091WJ2	091WJ2 mus musculu
4	601	42.7	553	4 096S37	096S37 homo sapien
5	592	42.1	550	4 09NSA0	09NSA0 homo sapien
6	551	39.2	553	11 054778	054778 mus musculu
7	458.5	32.6	506	4 09NQ02	09NQ02 homo sapien
8	458.5	32.6	519	4 09NQ06	09NQ06 homo sapien
9	458.5	32.6	563	4 09S742	09S742 homo sapien
10	457.5	32.5	551	11 035956	035956 ratu
11	449.5	32.0	551	6 09TSY7	09TSY7 oryctolagus
12	439	31.2	562	13 057379	057379 pseudop
13	438.5	31.2	545	11 061185	061185 mus musculu
14	390.5	27.8	542	4 096T01	096T01 homo sapien
15	372.5	26.5	552	4 096DT2	096DT2 homo sapien
16	368.5	26.2	537	11 091WJ9	091WJ9 mus musculu

17	367.5	26.1	537	11 088909	088909 mus musculu
18	362.5	25.8	536	11 09R107	09R107 ratu
19	324.5	23.1	535	11 063314	063314 ratu
20	318	22.6	540	11 091WJ2	091WJ2 mus musculu
21	312.5	22.2	140	11 09QYX0	09QYX0 ratu
22	307.5	21.9	539	4 09H2M5	09H2M5 homo sapien
23	306	21.8	568	4 09S820	09S820 homo sapien
24	296.5	21.1	548	4 09Y694	09Y694 homo sapien
25	234	16.6	551	4 09Y226	09Y226 homo sapien
26	228	16.2	538	5 09V6H5	09V6H5 drosophila
27	228	16.2	604	5 0961J5	0961J5 drosophila
28	223	15.9	548	5 001384	001384 drosophila
29	223	15.9	348	5 09VCA2	09VCA2 drosophila
30	220	15.6	564	11 09WTN6	09WTN6 mus musculu
31	202.5	14.4	553	11 09R141	09R141 ratu
32	197.5	14.0	553	11 09Z306	09Z306 mus musculu
33	197.5	14.0	568	5 09U539	09U539 caenorhabd
34	197.5	14.0	576	5 002270	002270 caenorhabd
35	192.5	13.7	567	5 095R48	095R48 drosophila
36	192	13.7	370	5 001383	001383 drosophila
37	191.5	13.6	567	5 09VCA3	09VCA3 drosophila
38	190.5	13.5	555	5 0961R9	0961R9 homo sapien
39	190.5	13.5	557	4 096EH6	096EH6 homo sapien
40	189.5	13.5	577	4 096RU0	096RU0 homo sapien
41	184	13.1	177	4 096ER0	096ER0 homo sapien
42	182.5	13.0	502	5 09V8Z3	09V8Z3 drosophila
43	179.5	12.8	551	4 09H015	09H015 homo sapien
44	179.5	12.8	557	5 09VEX8	09VEX8 drosophila
45	178.5	12.7	551	4 014546	014546 homo sapien

ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	552 AA.
070609	070609			
AC	01-AUG-1998 (TREMBLrel. 07, Created)			
DT	01-AUG-1998 (TREMBLrel. 07, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	PUTATIVE INTEGRAL MEMBRANE TRANSPORT PROTEIN.			
GN	USP1R.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=KIDNEY;			
RX	MEDLINE=98200080; PubMed=9541011;			
RA	Schoenly E., Spitzberger F., Engelhardt M., Martel F., Oeding N.,			
RA	Gruendemann D.,			
RT	"Molecular cloning and characterization of two novel transport			
RT	proteins from rat kidney."			
RL	FEBS Lett. 425:79-86(1998).			
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).			
CC	-1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.			
DR	EMBL: Y09945; CA71076.1; -			
DR	InterPro: IPR003662; sub_transporter.			
DR	Pfam: PF00083; sugar_tr 1.			
KW	Transmembrane.			
SQ	SEQUENCE 552 AA; 61476 MW; F47FBA2B41F5C4AF CRC64;			

Query Match: 58.2%; Score 818; DB 11; Length 552;
Best Local Similarity: 45.9%; Pred. No. 9.5e-71;
Matches 164; Conservative 40; Mismatches 45; Indels 108; Gaps 1;

QY 1 MAFFELLISVGLGRFQMLHVLPIISLMLLPHILLNFPAALPGHRCVHMLDNTGS 60
DB 1 MAFFOLLNVOGSLGRFQMLHVLPIISLMLLPHILLNFPAALPGHRCVHMLDNTGS 60